②

主論文

# 学位申請論文

大腸菌における GC-CG transversion 型突然変異
の抑制機構の研究

中原岳久

# 要旨

塩基置換型の突然変異のうち、GC 塩基対から CG 塩基対への突然変異（GC-CG　トランスバージョン型）の自然発生頻度はその他のトランスバージョン、トランジションに比べて低いことが知られている。しかし、このタイプの突然変異は活性酸素やUV-B、γ線などで処理すると発生頻度が上昇する。このことは、GC-CG　トランスバージョンを起こす原因となる塩基の損傷（修飾）が存在すること、この修飾塩基はDNA上に生成しても通常は修復されることを示している。ところが、このタイプの損傷を修復するシステムに欠損の生じた細胞、つまりこの突然変異を特異的に上昇させるミューテーター株は見つかっていない。GC-CG トランスバージョンを抑制する機構の解明を目的として、二つの方向からのアプローチを行った。一つはこの突然変異が生じる中間段階のミスマッチを特異的に認識するタンパクを探すことであり、もう一つはこれまで分離されていなかった大腸菌のミューテーター変異株を分離することである。まず、ゲルシフト法を用いて大腸菌中でのミスマッチ認識タンパクを探したところ、シトシン（C）を含むミスマッチ、特に C-C のミスマッチを強く認識する活性が2つ検出された。うち一つは、様々な変異株およびクローニングした遺伝子を用いた実験の結果、MutM タンパクであることが分かった。酵母（*Saccaromyces cerevisiae*）の Ogg1 タンパクも同様に C-C ミスマッチに結合することが分かった。一方、もう一つのタンパクを精製し、その N 末端のアミノ酸配列から大腸菌 ORF データベースを検索したところ、FabA タンパクであることが分かった。これらのタンパクは DNA の C-C ミスマッチの修復に関与しているものと考えられる。

次に、トランスポゾンをランダムにゲノム中に組み込む方法を用いて、GC-CG トランスバージョンのミューテーター株を分離することができた。このトランスポゾンの末端配列を利用して変異している遺伝子の配列を調べたところ、*mutY* 遺伝子であることが分かった。精製した MutY タンパクは、8-オキソグアニンとアデニンのミスマッチだけでなく、8-オキソグアニンとグアニンのミスマッチも認識して修復することが分かった。

目次

# 序論

突然変異はまれにしか起こらない現象である。しかし、その生物学的な意義は大きい。プラスの面では、長い進化の過程において、様々な突然変異の蓄積と選択によって現在のような多様な生物種を生み出してきたという点があげられる。すなわち、突然変異は進化の原動力であるということができる。しかし、個々の個体にとっては、ほとんどの突然変異は役に立たないか、有害であると考えられる。このように、突然変異は起こりすぎても種の維持に困難を来す。そのため、生物は突然変異を抑制する機構を発達させてきた。このようなシステムを、複雑さや構成因子の数などの違いはあるが、ウイルスからヒトまでが持っている（Friedberg *et al.* 1995）。

点突然変異の原因としては、DNA 複製時のエラーや損傷塩基による誤った対合の形成、相同だが完全には一致しない配列間での組み換えなどで生じるミスマッチがあげられる（Schaaper *et al.* 1987, Dra1991, Echols *et al.* 1991, Modrich 1991, Smith 1992, Miller 1992, Modrich *et al.*1996, Wallace 1997）。このようなミスマッチを修復することは突然変異の抑制やゲノムの安定化にとって重要である（Miller 1992, 1996, 1998, Modrich *et al.* 1996, Wallace 1997, Mol 1999, Eisen 1999）。現在まで、様々な生物種で、点突然変異を抑制する機構に関する研究が行われてきた（Modrich 1991, Modrich *et al.*1996, Mol *et al.*1999, Eisen *et al.* 1999）。点突然変異の抑制には DNA ポリメラーゼの校正機能やミスマッチ修復系が関与していることが明らかにされてきた（Schaaper, 1993）。DNA ポリメラーゼ自身の突然変異頻度はおよそ $10^{-4}$ /塩基対/複製と見積もられている。これをポリメラーゼの校正機能やミスマッチ修復系などを用いて$10^{-10}$/塩基対/複製まで減少させている（Schaaper, 1993）。ミスマッチ修復系は大腸菌からヒトまで、生物種を越えて基本構成が保存されている DNA 修復系である（Modrich 1991, Modrich *et al.* 1996, Miller 1996, 1998, Mol *et al.* 1999, Eisen *et al.*

1999）。主な構成タンパクは、G-Tのようなミスマッチを認識するMutS Mshタンパクと、機能の詳細は明らかになっていないがMutS/Mshタンパクと結合するMutL/Mlh/Pmsタンパクである。また、大腸菌ではGATC配列のアデニンのメチル化を指標として鋳型鎖と新生鎖を識別し、新生鎖のみこの配列部分を切断するMutHタンパクも同定されている（Modrich 1991, Miller 1992, 1996, 1998, Modrich *et al.* 1996）。これらミスマッチ修復系の遺伝子の機能が失われると、自然突然変異頻度が上昇する（Miller 1992, 1996, 1998, Modrich *et al.* 1996）。

大腸菌では、早くから簡便に点突然変異を検出するシステムが開発されてきた。その代表的なものがMillerらが開発したCCシリーズと呼ばれるものである（Cabrera *et al.* 1988, Nghiem *et al.* 1988, Cupples *et al.* 1989）。大腸菌のF'因子上のβーガラクトシダーゼ遺伝子中に一カ所だけ、六つの塩基置換型変異のうち一つの変異が導入されている（Cupples *et al.* 1989）。この変異が野生型に復帰する現象を観察することで突然変異の頻度や傾向を知ることができる。

このシステムなどを用いて様々な変異株の突然変異の自然発生頻度が調べられてきたが、CC103株の突然変異頻度を特異的に上昇させる大腸菌の変異株は発見されなかった（Miller 1992, 1998）。CC103株はGCからCGへの突然変異が生じたときに野生型に復帰する株である。このことは、これまで述べてきた様々なDNA修復機構はGCからCGへのトランスバージョンを抑制しないということを意味している。しかし、突然変異を誘発する処理、たとえば一重項酸素やUV-B、γ線などの処理によってこのトランスバージョンの頻度が上昇する（MacBride *et al.* 1991, Akasaka *et al.* 1994, Sargentini *et al.* 1994, Buchko *et al.* 1995）。また、*mutHLS*ミスマッチ修復機構は、GCからCGへの突然変異の中間段階で発生することが予想されるミスマッチのうち、G-Gのミスマッチを効率よく修復するが、C-Cのミスマッチは修復効率が低いことが知られている（Kramer *et al.* 1984, Su *et al.* 1988, Schaaper 1993）。この原因は、ミスマッチを認識するMutSタンパクがほとんどC-Cのミス

マッチを認識しないことだと考えられている（Wagner *et al.* 1995, Babic *et al.* 1996）。これらの現象から、GC から CG への突然変異を抑制している機構が存在すると予想される。私は、その機構を解明するべく本研究を行った。

本研究では、２つの方向からこの機構の解明を試みた。１つは突然変異の中間段階で形成される、何らかの形での C-C や G-G のミスマッチを認識するタンパクが存在するかどうかの確認である。もう一つは CC103 株で自然発生突然変異頻度を上昇させる変異株（ミューテーター株）を分離することである。

まず、C-C ミスマッチを認識するタンパクが存在するかどうかを大腸菌の粗抽出液を用いたゲルシフト法で検討した。この方法によってすでに分裂酵母の Cmb1 が発見されている（Fleck *et al.* 1994）。この実験の結果、C-C を主に認識するタンパクが存在することが確認できた。また、このタンパクは C-C のほかに C-A や C-T といったシトシンを含むミスマッチを認識することが分かった。このタンパクを同定するため、様々な変異株から得られた粗抽出液を用いたゲルシフト法で検討した。その結果、ゲルシフト法で検出される二本のバンドのうち、よりシフトする方が *mutM* 欠損株で無くなることが分かった。さらに、精製した GST-MutM タンパクを用いた実験の結果、MutM タンパク自身が C-C のミスマッチに結合することが分かった。しかし、この方法ではもう一方のバンドを形成するタンパクは同定できなかった。そこで、このタンパクを精製し、そのアミノ酸配列からタンパクを同定した。アミノ末端のアミノ酸配列を調べた結果、Val-Asp-Lys-Arg-Glu-?-Tyr-Thr-Lys-Glu であった。これを国立遺伝学研究所の大腸菌 ORF データベースで調べた結果、FabA タンパクのアミノ末端の配列と一致した。また、大腸菌 *fabA* 変異株から調整した粗抽出液を用いたゲルシフトアッセイでは、MutM とは違う方のバンドが消失することも分かった。

次に、GC から CG へのトランスバージョンのミューテーター株を作成した。トランスポゾンをランダムに大腸菌 CC103 株のゲノムに挿入する操作を行った。約

4万個のコロニーをP-GalとX-Galを含んだグルコース最小培地プレートにレプリカし、親株であるCC103株よりパピーレを明らかに多く生じるコロニーを選択した。選択した株からP1 lysateを調整し、もう一度CC103株に感染させる事でゲノム内に一つのトランスポゾンのみを含む株にし、なおかつ、そのトランスポゾンの挿入によってGCからCGへのトランスバージョンの発生頻度が上昇することを確認した。この結果、ミューテーター株を五つ得ることができた。

五つのミューテーター株の突然変異スペクトルをみたところ、GC39株だけがGCからCGへのトランスバージョンのみを上昇させることが分かった。それ以外の株はすべての種類の塩基置換型の突然変異が反映されるリファンピシン抵抗性への突然変異頻度がGCからCGへのトランスバージョンの発生頻度を大きく上回っている。そこで、本研究では、GCからCGへのトランスバージョンの発生頻度のみを上昇させるGC39株について詳しく調べることにした。この株からゲノムDNAを回収し、トランスポゾン内部の抗生物質耐性遺伝子を指標にして、この遺伝子を含む周辺配列をクローニングした。トランスポゾンの末端配列を利用してその境界部分の塩基配列を調べたところ、*mutY*遺伝子であることが分かった。そこで、MutYタンパクを精製し、GCからCGへのトランスバージョンの抑制との関係を生化学的に調べてみた。ゲルシフト法による検討の結果、MutYタンパクは今まで知られていた8−オキソグアニンとアデニンの塩基対だけでなく、8−オキソグアニンとグアニンの塩基対も認識することが分かった。さらに、MutYタンパクは8−オキソグアニンとアデニンの塩基対からアデニンを除去する効率の約1/7の効率であるが、明らかに8−オキソグアニンとグアニンの塩基対からグアニンを除去することが分かった。

## 材料と方法

基本的な分子生物学的実験は、プロトコール集であるMolecular Cloning（Fritsch *et al.* 1989）に、大腸菌の基本的操作は同じくプロトコール集であるA Short Course in Bacterial Genetics （Miller 1992） に従った。

### １、本研究に用いた大腸菌株

表１に本研究に用いた大腸菌株を記した。

### ２、本研究に用いた合成DNA

表２に本研究に用いた合成DNAを記した。

### ３、出芽酵母 *Ogg1* 遺伝子のサブクローニング

基本的な操作はGirardらの方法（Girard *et al.* 1997） に従った。放射線医学総合研究所の味村正博博士より分与していただいた出芽酵母（S. cerevisiae） ゲノムライブラリーを鋳型とし、表２の９番のプローブと１０番のプローブをプライマーとしてPCRを行い、出芽酵母の *Ogg1* 遺伝子を増幅した。用いた酵素はKOD Plus DNAポリメラーゼ（TOYOBO）である。増幅された1.1 kbの断片を制限酵素処理後プラスミドpKK223-3（Amersham Pharmacia Biotech）にクローニングした。このプラスミド（pKK223-3-YOGG1）を大腸菌CC104 *mutMmutY* 株に導入し、自然突然変異頻度を野生株レベルにまで低下させることを確認した。

## 4、大腸菌 *mutY* 遺伝子のサブクローニング

大腸菌 *mutY* 遺伝子を含むプラスミド pLC20-5（Nishimura *et al.* 1992）を鋳型とし、表2の12番のプローブと13番のプローブをプライマーとして PCR を行い、大腸菌 *mutY* 遺伝子を増幅した。この PCR 断片を *Eco*RI と *Sal*I で制限酵素処理を行い、*Eco*RI と *Sal*I で制限酵素処理を行ったプラスミド pGEX-4T-3 にクローニングした。得られたプラスミド（pGEX-MutY）を大腸菌 CC104 *mutMmutY* 株に導入し、自然突然変異頻度を野生株レベルにまで低下させることを確認した。

## 5、大腸菌粗抽出液の調整

大腸菌の終夜培養菌液を 6000 rpm、15分の遠心操作にかけて、菌を回収した。沈殿 1g あたり 3 ml のbufferA（50 mM Tri-HCl, pH 8.0、1 mM EDTA）に溶かし、-80℃で保存した。このサンプルをさらに bufferA で3倍希釈し、氷水中で超音波処理を行った。この後、8000 rpm と 12000 rpm で4゜C、30分ずつ遠心操作にかけ、その上清を粗抽出液として、以後の実験に用いた。

## 6、出芽酵母粗抽出液の調整

出芽酵母の培養菌液を 6000 rpm、5分の遠心操作にかけて、菌を回収した。さらに冷やした滅菌蒸留水で菌を洗った。この菌に抽出バッファー（20 mM Tris-HCl, pH 7.5、5 mM $MgCl_2$、1 μg/ml pepstatin A、1 mM PMSF、3 mM dithiothreitol (DTT)）を加え、ボルテックスミキサーを使って菌を溶かした。ここにガラスビーズ（直径 0.5 mm）を加えさらにボルテックスミキサーを使って攪拌した。攪拌終了後、飽和硫酸アンモニウム溶液を最終濃度 10 %になるように加え、ボルテックスミキサー

を使ってゆるやかに攪拌した。この溶液を 14600 rpm で３０分遠心操作にかけ、その上清を粗抽出液として、以後の実験に用いた。

## ７、ゲルシフトアッセイ

基本的な操作は Fleck らの方法（Fleck *et al.* 1994）に従った。合成 DNA をＴ４ポリヌクレオチドキナーゼと [γ-$^{32}$P] ATP を用いて標識し、相補鎖とアニーリングさせたものをプローブとして用いた。反応液の組成は 25 mM Tri-HCl（pH 8.0）、0.5 mM DTT、4 mM spermidine、0.5 mM EDTA、10 % glycerol、25 mM NaCl、25 mM KCl、0.01 mM $ZnCl_2$、0.125 mM の dATP、dCTP、dGTP、dTTP、そして 5 mg の超音波処理済みの子牛胸腺 DNA という条件で行った。二重鎖のオリゴヌクレオチドと粗抽出液を混合し、4° C、３０分反応後、12 %の未変性ポリアクリルアミドゲルで電気泳動を行った。泳動後、ゲルを乾燥させてＸ線フィルム（Fuji RX）に-80 ° Cで感光させた。

## ８、GST 融合タンパクの精製

GST-MutM タンパクは、国立ガンセンターの横田淳博士から分与していただいたプラスミド（Nagashima *et al.* 1997, Shinmura *et al.* 1997）を、大腸菌 BL21 株に導入した株から回収した。この株の終夜培養液を 100 倍希釈し、37C、２時間培養後、最終濃度 1 mM の IPTG を加え、さらに 37 ° C、８時間培養した。また、GST-MutY タンパクはプラスミド pGEX-MutY を大腸菌 BL21 株に導入した株から回収した。この株を 22-24° Cで培養し、$OD_{600}$ が 0.8 になったときに最終濃度 0.1 mM の IPTG を加え、さらに 22-24 ° Cで８時間培養した。これらの培養液から遠心操作によって大

腸菌を回収し、超音波処理後、遠心操作を行ってその上清を回収した。その上清から Glutathione-Sepharose 4B カラム（Amersham Pharmacia Biotech）を用いて GST 融合タンパクを回収した。操作などはカラムの取扱説明書に従った。回収した GST 融合タンパクはトロンビンプロテアーゼ処理によって GST-tag をはずし、以後の実験に用いた。

## ９、C-C ミスマッチ結合タンパクの精製

大腸菌 *mutM* 欠損株から調整した粗抽出液を硫酸アンモニウム沈殿法によって 40-60 %で沈殿する分画を回収した。この分画を bufferA に溶かした後、bufferA を外液として透析した（分画１）。分画１を HiTrap Q 陰イオン交換カラム（Amersham Pharmacia Biotech）に通し、bufferB（20 mM Tri-HCl, pH 8.0、150 mM NaCl）でカラムを洗浄したときにカラムに結合しない分画を回収した（分画２）。分画２を bufferC（20 mM Hepes, pH 7.6、150 mM NaCl）で平衡化した HiTrap-Heparin カラム（Amersham Pharmacia Biotech）で分画した（分画３）。さらに、分画３を HiPrep 16/60 Sephacryl S-200HR ゲル濾過カラム（Amersham Pharmacia Biotech）にかけ、bufferB で溶出した（分画４）。すべての段階でゲルシフト法による確認を行って、活性のある分画を回収した。

## １０、C-C ミスマッチ結合タンパクのアミノ酸配列の決定

精製した分画４を17.5 %の SDS-PAGE によって分離し、セミドライエレクトロブロッティング法で分離したタンパクを Immobilon-P$^{SQ}$ メンブレン（Millipore）にブロッティングした。このメンブレンをクマシーブルーで染色し、約 20 kDa のタンパクのバンドを切り取り、気層アミノ酸シークエンサーによってアミノ末端の１０

アミノ酸残基を決定した。この配列を元に国立遺伝学研究所の大腸菌 ORF データベースで検索した。

## １１、ミューテーター株の探索

ミューテーター株の作成には Randam Tn *10* insertion mutagenesis 法を用いた。方法の詳細は A Short Course in Bacterial Genetics（Miller 1992）に従ったので、概略のみを記す。

λファージ λ1098（Kleckner *et al.* 1978）を CC103 株に感染させ、miniTn *10*-tet トランスポゾンを大腸菌ゲノム内にランダムに挿入した。これを lym プレート（Kleckner *et al.* 1978）にプレーティングし、39.5℃で２０時間ほど培養した。約4万個のコロニーを 15 μg/ml のテトラサイクリン、40 μg/ml の 5-Bromo-4-chloro-3-indolyl-D-galactside（X-Gal）と 500 μg/ml の Phenyl- β-D-galactside（P-Gal）を含むグルコース最小培地にレプリカした。これらのコロニーのうち明らかにパピーレの出現頻度の高いものを選抜し、P1 lysate を調整した。この lysate を用いた P1 トランスダクション法でもう一度 CC103 株にトランスポゾンを移し、各変異株にトランスポゾンが1個のみ入っているようにした。

## １２、自然突然変異頻度の算出

各変異株で single colony isolation を行い、得られたコロニーをグルコース最小培地中、37℃で終夜培養した。この菌液 0.1 ml をラクトース最小培地にプレーティングし、37℃で１０日間培養した。また、この菌液 0.1 ml を 100 μg/ml のリファンピシンを含む LB プレートにプレーティングし、37℃で２４時間培養した。培養後、生育してきたコロニー数を数え、自然突然変異頻度を求めた。

## １３、ミューテーター遺伝子のマッピング

Singer の P1 マッピングキットの大腸菌（Singer *et al.* 1989）から P1 lysate を調整し、GC39 株に形質導入した。カナマイシンやクロラムフェニコール耐性コロニーをテトラサイクリンを含む LB プレートにレプリカし、それらのうちでテトラサイクリン感受性のコロニーの割合を求めた（Berlyn *et al.* 1996, Singer *et al.* 1996）。カナマイシンもしくはクロラムフェニコール耐性でかつテトラサイクリン感受性のコロニーを P-Gal と X-Gal を含むグルコース最小培地にレプリカし、パピーレの出方を観察した。

## １４、*mut*39::miniTn*10*遺伝子のサブクローニング

大腸菌 GC39 株のゲノム DNA を回収し、*Bam*HI 処理後、pUC118 の *Bam* HI サイトにサブクローニングした。このプラスミドを大腸菌 JM109 株に導入し、テトラサイクリンとアンピシリンを含む LB プレートにプレーティングした。ここで生えてきた菌からプラスミドを回収し、*mut39* と miniTn*10* の境界部分を BcaBest™ シークエンスキットと miniTn *10* の末端配列（Halling et al. 1982）を利用したプライマー（表２の１１番）を利用してシークエンスした。電気泳動は６％変性アクリルアミドゲルで行い、ゲルを乾燥後、-80°CでX線フィルムに感光させた。

## １５、ニッキングアッセイ

配列中に８－オキソグアニンを１カ所で含む標識２本鎖 DNA 20 fmol に様々な量の MutY タンパクを加え、10 ml の 20 mM Tris-HCl（pH 7.6）、0.5 mg BSA、10

mM EDTA 中で 37 ℃で反応させた。反応後、2 ml の 2.5 M NaCl を加え、95°Cで５分インキュベートした。その後、95 %ホルムアミド、0.1 %ブロモフェノールブルー、0.1 %キシレンシアノール、20 mM EDTA 溶液を加え、更に 98 °Cで５分インキュベートした。このサンプルを冷却し、7 M 尿素を含む 20 %アクリルアミドゲルで電気泳動を行った。ゲルを乾燥後、-80 °CでＸ線フィルムに感光させた。

# 結果

## １－１、C-C や G-G のミスマッチを認識するタンパクの検出

GC-CG のトランスバージョンは、その中間段階で何らかの形での C-C や G-G ミスマッチを経ることが予想される。しかし、これまでのところ、GC-TA の突然変異における８－オキソグアニンのような、GC-CG の突然変異を誘発すると思われる損傷塩基は５－ハイドロキシシトシンが報告されているが、未だ確実とは言い難い（Purmal et al. 1994a, b）。初期の大腸菌 MutY タンパクの活性検出法は、現在知られている８－オキソグアニンとアデニンの塩基対からではなく、単なるグアニンとアデニンのミスマッチからアデニンを切り出させるという方法であった（Aboul-ela *et al.* 1985）。このことから、損傷塩基を含むミスマッチを認識する活性を正常な塩基同士のミスマッチでも検出できるかもしれない考え、様々なミスマッチ、特に C-C や G-G を認識するタンパクが大腸菌野生株中に存在するかどうかをゲルシフト法を用いて検討した（図１）。この実験の結果、レーン１の C-C のミスマッチに２本のバンドが確認された。G-C やそれ以外のミスマッチにはほとんど結合しないことも分かった。別の配列を用いた同様の実験でも、C-C のミスマッチに２本のバンドが確認された（data not shown）。

さらに、C-C のプローブに対するコンペティションアッセイを行った（図２）。コンペティターの量はプローブの 50 倍を加えた。この条件では、レーン１の C-C のコンペティターを加えたレーンのみ、２本ともバンドが消失している。これらのことから、大腸菌には、C-C に結合するタンパクが１つ以上存在することが分かった。今回の実験条件では、G-G に結合するタンパクは検出できなかった。

C-C 以外にも弱いながら、アフィニティーのある組み合わせがある。C-C 以外のシトシンを含むミスマッチ、C-T や C-A などである。これらを比較してみた（図

３）。プローブはレーンの５から８がC-C、９から１２がC-A、１３から１６がC-Tである。アフィニティーの強さはC-Cが最も強く、次にC-T、C-Aの順番になっている。しかし、レーンの１０や１４でC-Cのコンペティターを加えるとバンドが消失することから、C-Cに比べて、かなりアフィニティーが弱く、これらの活性の主要な基質ではないだろうと思われる。

１－２、様々な変異株中の C-C ミスマッチ結合タンパクの有無の確認

次に、これらのC-C結合タンパクが今までに確認されているミューテーター遺伝子などの産物なのかどうかを検討してみた。

様々な変異株から得られた粗抽出液を使ってゲルシフトアッセイを行った（図４）。その結果、レーン３の *mutM* 以外の変異株には２本ともバンドがあり、これらの遺伝子産物は今回解析しているものではないことが分かった。また、レーン３の *mutM* 欠損株で、２本あるバンドのうち、バンド１だけが消失した。このことから、この２本のバンドは同じ種類のタンパクが２つ結合したものではなく、異なるタンパクが結合していることが予想された。

そこで、MutMタンパクが直接このバンド１を形成しているのかどうかを調べるため、MutMタンパクを精製してゲルシフトアッセイを行った。発現系として、GST-fusionのシステムを用いた（Nagashima *et al.* 1997, Shinmura *et al.* 1997）。

その結果、GST-tagを除去したMutMタンパクが *mutM* 変異株で消失するバンド１とほぼ同じ位置にバンドを形成することが分かった（図６）。また、その結合の特異性も同じであることが分かった（図７）。

以上のことから、このバンド１はMutMタンパク自身が結合したものであることが分かった。しかし、もう一つのタンパクはこの実験では明らかにできなかった。

## １－３、もう一つの C-C ミスマッチ結合タンパクの精製

また、今回の研究で用いた様々な変異株でも結合活性の残っている、バンド２に関して、コードしている遺伝子や生化学的な性質など、詳しいことを明らかにするため、このタンパクの精製を行った。

それぞれのステップでのサンプルを SDS-PAGE で確認したものが図８である。このように、最後のサンプルはほぼ単一のバンドになっている。このバンドからタンパクを PVDF メンブレンにブロットし、N 末端のアミノ酸配列の解析を行った。

N 末端の配列を１０残基分調べたところ、VDKRE?YTKE というようになった（図９）。? は解析できなかったものである。この配列を元に、遺伝学研究所の大腸菌の ORF データベースを検索したところ、大腸菌の FabA タンパクの N 末端のアミノ酸配列と一致することが分かった。

そこで、大腸菌 *fabA* 温度感受性変異株から粗抽出液を調整し、ゲルシフト法で確認してみた（図９）。37 ℃で培養したところ、図９のレーン２のように FabA のバンドが消失することが分かった。

## １－４、MutM タンパクの機能ホモログでの C-C ミスマッチ認識能の確認

大腸菌 MutM タンパクの機能ホモログは出芽酵母やヒトなどで発見されている（van der Kemp *et al.* 1996, Nash *et al.* 1996, Sandigursky *et al.* 1997, Bjoras *et al.* 1997, Dherin *et al.* 1999, Ishida *et al.* 1999）。しかし、これらのタンパクと大腸菌 MutM タンパクとの間にはこれといったホモロジーは無い（van der Kemp *et al.* 1996, Nash *et al.* 1996）。これらのタンパクに大腸菌 MutM と同じように C-C のミスマッチを認識する能力があるかどうかを確認してみた（図１０、１１）。出芽酵母の *Ogg1* やヒ

トの　*Ogg1*のいくつかのタイプをクローニングしたプラスミドを導入した大腸菌*mutM*欠損株の粗抽出液を用いたゲルシフト法で調べた。その結果、出芽酵母のOGG1、およびヒトのOGG1の１つのタイプがC-Cのミスマッチに結合することが分かった。

図１０で示したように、出芽酵母のOGG1タンパクは大腸菌MutMとほぼ同じ選択制を示したが、図１１で示したヒトのOGG1は結合能が弱いためもしくは発現量が少ないためか、C-Cのミスマッチに対する結合しか確認できなかった。また、出芽酵母野生株から得られた粗抽出液を用いたゲルシフトアッセイも行ってみた（図12）。すると、OGG1タンパクとは異なる位置にかなり強いバンドを形成することが分かった。このバンドのミスマッチ認識の特異性は、OGG1タンパクやMutMとほぼ同じでシトシンを含むミスマッチを主に認識している。さらに、アデニン同士のミスマッチも認識できることも分かった。

２－１、ミューテーター株の探索

次に、大腸菌CC103株（Cupples *et al.* 1989）（図１３）に直接トランスポゾンを導入することで、ランダムに遺伝子を破壊した菌株を作成し、その中からGC-CGトランスバージョンの自然突然変異頻度が高くなるような菌株（ミューテーター株）を探索した。現在までに大腸菌で見つかっている様々なミューテーター株には、CC103株の自然突然変異頻度だけを上昇させるようなものは見つかっていない。約４万個のトランスポゾン挿入コロニーを調べた結果、５つの候補が得られた（図１４）。これらをさらに調べた結果、３つのグループに分けることができた（表３）。グループ１はパピーレを多く発生させ、$Lac^{+}$への復帰突然変異とリファンピシン耐性の突然変異の数がほぼ同じものである（GC39株）。グループ２はパピーレを多く発生させるが、$Lac^{+}$への復帰突然変異とリファンピシン耐性の突然変異の

数が大きく違うものである（GC9、GC24 株）。グループ3はパピーレをあまり多く発生させず、$Lac^+$ への復帰突然変異とリファンピシン耐性の突然変異の数が大きく違うものである（GC40、GC41 株）。これらのうち、グループ2と3はリファンピシン耐性突然変異が $Lac^+$ への復帰突然変異よりもかなり多く発生することから、GC から CG への突然変異以外の突然変異が多く起こっているものと考えられる。そのため、純粋に GC から CG への突然変異のみが多くなっていると思われる GC39 株を解析することにした。

## ２－２、GC39　株におけるミューテーター遺伝子のマッピング

Singer の P1 マッピングキットの大腸菌（Singer *et al.* 1989）から P1 ファージの lysate を調整し、GC39 株に感染させた。得られたカナマイシンやクロラムフェニコール耐性コロニーをテトラサイクリンを含む LB プレートにレプリカし、それらのうちでテトラサイクリン感受性のコロニーの割合を求めた。この方法（Berlyn *et al.* 1996, Singer *et al.* 1996）は、カナマイシンやクロラムフェニコール耐性とテトラサイクリン感受性が同時に成立する割合が高ければ高いほど、GC39 株のミューテーター遺伝子（*mut39*）と薬剤耐性遺伝子が同時に形質導入される、すなわち両遺伝子間の距離が近いという性質を利用したものである。この実験の結果、大腸菌ゲノム上の 66.9 分のトランスポゾンで 90%、67.9 分で 23 %、66.0 分で 19 %のコロニーがテトラサイクリン感受性でカナマイシン耐性であった。このことから、*mut39* 遺伝子は大腸菌ゲノム上の 67 分付近に存在することが分かった。

## ２－３、*mut39*::miniTn*10* 遺伝子のサブクローニング

次に、*mut39* 遺伝子内に挿入されている miniTn *10* トランスポゾンの末端の配

列（Halling *et al.* 1982）を利用して、挿入された部分の周囲の塩基配列を決定することにした。サブクローニングした *mut39*::miniTn*10* 遺伝子をシークエンスした結果を遺伝学研究所の大腸菌データベースで検索した結果、大腸菌の *mutY* 遺伝子の配列であることが分かった（Michaels *et al.* 1990a, Tsai-Wu *et al.* 1991）。また、GC39 株に *mutY*$^{+}$ をコードする遺伝子をのせたプラスミドを導入したところ、GC から CG への変異頻度が大きく減少した（表４）。さらに、*mut39*::miniTn*10* 遺伝子を CC104 株（Cupples *et al.* 1989）に導入すると *mutY* 変異株（Nghiem *et al.* 1988）とほぼ同じくらいの突然変異頻度を示した。以上のことから、*mut39* 遺伝子は *mutY* 遺伝子であることが分かった。

２－４、MutY タンパクを用いたゲルシフトアッセイ

GC39 株で変異している、GC-CG の突然変異を誘発するミューテーター遺伝子が *mutY* 遺伝子であることが分かったので、プラスミド pGEX-4T-3（Amersham Pharmacia Biotech）に *mutY* 遺伝子をクローニングしたものを構築し、GST-MutY タンパクを精製した。トロンビンプロテアーゼ処理によって GST-tag をはずした MutY タンパクを以下の実験に用いた。

MutY タンパクは、８－オキソグアニンとアデニンの塩基対からアデニンを除去する酵素として知られている。MutY タンパクが８－オキソグアニンとグアニンの塩基対を認識することができるかを調べるため、ゲルシフトアッセイを行った（図１５）。その結果、MutY タンパクの量にほぼ比例して、８－オキソグアニンとアデニンの塩基対だけでなく８－オキソグアニンとグアニンの塩基対も認識できることが分かった。また、８－オキソグアニンを含まない塩基対は殆ど認識しないことも分かった。また、その結合効率はほぼ同じであることが分かった（data not shown）。

## ２－５、ニッキングアッセイ

さらに、MutY タンパクが8－オキソグアニンとグアニンの塩基対からどちらかの塩基を除去することができるのかどうかを調べるために、ニッキングアッセイを行った（図１６）。プローブはゲルシフトアッセイと同じものを利用した。その実験の結果、８－オキソグアニンとグアニンの塩基対の場合、グアニンが除去されることが分かった。これは、８－オキソグアニンとアデニンの塩基対からアデニンが除去されるのと同じ選択性を示している。ただし、その除去の反応速度は８－オキソグアニンとアデニンの塩基対の場合よりも遅くなっている（図１７）。８－オキソグアニンとグアニンの塩基対では約8時間でほぼ完全に切断されているが、同じ条件下では ８－オキソグアニンとアデニンの塩基対はほぼ1時間で完全に切断される。さらに、それぞれの塩基対に対する比活性を求めた（図１８）。その結果、８－オキソグアニンとアデニンの塩基対からアデニンを除去する活性は 0.7 nmol/mg、８－オキソグアニンとグアニンの塩基対からグアニンを除去する活性は 0.1 nmol/mg であった。

## 考察

### 大腸菌野生株でのミスマッチ認識活性の検出

DNA 複製時のエラーはC-CやG-Gのようなミスマッチ生成の潜在的な一要因である（Modrich 1996, Miller 1996, Miller 1998）。相同であるが同一ではないDNA配列間での組換えもミスマッチの要因である（Fleck *et al.* 1994, Fox *et al.* 1994, Worth *et al.* 1994）。DNA中のミスマッチはミスマッチ修復の基質となる。C-Cのミスマッチはその他のタイプのミスマッチに比べて不安定なので（Aboul-ela *et al.* 1985）、ミスマッチ修復のよい基質であると考えられる。しかし、*mutHLS*ミスマッチ修復系はC-Cのミスマッチを認識しない（Miiler 1998）。もしC-Cのミスマッチが修復されなければ、GC-CGの突然変異の原因となるはずである。しかし、GC-CGの突然変異の自然発生頻度はきわめて低いことが知られている（Miiler 1998）。そのため、大腸菌ではC-Cのミスマッチに対する特異的な修復系が無ければならないと考えた。本研究では、大腸菌野生株の粗抽出液中にC-Cのミスマッチを含むオリゴヌクレオチドを認識する２つのタンパクが存在することを確認した（図１）。さらに、標識していないC-Cのミスマッチを含むオリゴヌクレオチドがC-AやC-Tのミスマッチに対する複合体の形成を阻害することも分かった（図２、３）。

### 大腸菌 MutM タンパクの機能の解析

本研究では、大腸菌MutMタンパクがC-Cミスマッチに特異的に結合することができることを示した（図４）。この結論は *mutM*欠損株の粗抽出液を用いたゲルシフトアッセイの結果、移動度のより低い方の複合体がなくなるという事実から導いたものである。さらに、精製したMutMタンパクがC-Cのミスマッチに直接結

合することができることも分かった（図５、６）。MutM タンパクは *in vitro* でホルムアミドピリミジン（FaPy）、８－オキソグアニン、５－ヒドロキシシトシン、チミングリコールなど幅広い損傷塩基を認識する（Hatahet *et al.* 1994, Tchou et al 1991, Tchou *et al.* 1994, David *et al.* 1998）。MutM タンパクはこれらの酸化的損傷塩基を除去する DNA グリコシラーゼ活性とその後の DNA を切断する AP リアーゼ活性を持つ（Mol *et al.* 1999, Eisen *et al.* 1999, Hatahet *et al.* 1994, Tchou et al 1991, Tchou *et al.* 1994, David *et al.* 1998）。本研究では精製した MutM タンパクは C-C のミスマッチを切断しないことが分かった（data not shown）。そのため、MutM タンパクは C-C ミスマッチを認識して結合することができるが、ミスマッチを修復するために DNA を切断する他の酵素が必要であることを示唆している。MutM タンパクはいくつかのタンパクで構成される C-C ミスマッチの修復機構の一要素だと考えている。

第２の C-C ミスマッチ認識タンパクの同定

さらに本研究では C-C のミスマッチに特異的に結合する第２のタンパクが大腸菌の FabA タンパクであることを示した（図７、８）。精製した FabA タンパクがミスマッチを含むオリゴヌクレオチドに結合してより移動度の高い方の複合体を形成することができる。さらに、*fabA* 欠損株の粗抽出液中にはこの結合活性が無い（図９）。FabA タンパクは大腸菌の脂肪酸合成系の酵素の一つであり、飽和脂肪酸に二重結合を導入する $\beta$-hydroxydecanoyl-acyl carrier protein dehydrase であることが知られている（Cronan *et al.* 1988, Silbert *et al.* 1967, Magnuson *et al.* 1993, Saito *et al.* 1995）。しかし、FabA タンパクのアミノ酸配列中にはDNA 結合ドメインは見つかっていない（Cronan *et al.* 1988）。おもしろいことに、*S. pombe* の脂肪酸合成酵素である p190/210 複合体（FAS）は１本鎖や２本鎖の DNA に結合して大きな DNA の凝縮体を形成させることができる（Kaslan *et al.* 1994）。さらに、１本鎖の相補 DNA の

renaturing 活性や strand exchange 活性を示す（Kaslan *et al.* 1994）。このように、脂肪酸合成系の酵素の一部は DNA-processing 活性も持っている。大腸菌での C-C ミスマッチの修復に MutM と FabA がどのように関わっているかを明らかにすることは今後の課題である。

大腸菌以外の生物における C-C ミスマッチ認識機構

*S. pombe* には、少なくとも２つのミスマッチ修復系が存在している。主要なシステムは C-C 以外のミスマッチを認識する大腸菌の *mutHLS* システムのようなものである。マイナーなシステムは C-C のミスマッチを認識し、 *mutHLS* システムよりもかなり短い DNA を除去する（Schar et al. 1993a, b）。さらに、 *S. pombe* には２つの異なった C-C ミスマッチ結合タンパク、Cmb1 と Cmb2 の存在が報告されている（Fleck *et al.* 1998）。精製した Cmb1 タンパクは 22 kDaの分子量である。予想されるアミノ酸配列から、このタンパクは DNA 結合モチーフの一つであるＨＭＧドメインを持っている。Cmb1 とは異なり、MutM と FabA は HMG ドメインを持っていない（Magnuson *et al.* 1993, Boiteux *et al.* 1987）。Cmb1 タンパクは C-A や C-T の他に、T-T、C/Δ のループ、Ｏ$^{6}$-メチルグアニンとシトシンのミスマッチ、シスプラチンによる GpG の分子内クロスリンクを認識することができる（Fleck *et al.* 1998）。それとは別に、Cmb2 は C-C のミスマッチしか認識しない（Fleck *et al.* 1998）。さらに、C-C のミスマッチの修復系にヌクレオチド修復系の酵素が関与していることも示唆されている（Fleck *et al.* 1999）。このことから、大腸菌でも *mutHLS* 系以外の DNA 修復系が C-C のミスマッチの修復を行っていることが予想される。MutM と FabA はそのような修復系の一要素で、MutS タンパクのようなミスマッチを認識するタンパクなのかもしれない。

## 真核生物における MutM ホモログの C-C ミスマッチ認識能力の確認

これまで、*S. cerevisiae*（Miret *et al.* 1993）やヒト（Jiricny *et al.* 1988）の粗抽出液中にはC-Cのミスマッチを特異的に認識する活性はほぼないと報告されている。*S. cerevisiae*、ラット、ヒトでは、OGG1タンパクが8－オキソグアニン-DNAグリコシラーゼ/ APリアーゼとして同定されている（van der Kemp *et al.* 1996, Nash *et al.* 1996, Sandigursky *et al.* 1997）。これらの酵素は大腸菌MutMタンパクと同じように８－オキソグアニンとシトシンやグアニンの塩基対から８－オキソグアニンを除去する（Mol *et al.* 1999, Eisen *et al.* 1999, van der Kemp *et al.* 1996, Nash *et al.* 1996, Sandigursky *et al.* 1997, Wallace 1997）。本研究で、*S. cerevisiae* とヒトのこれらの酵素がC-Cのミスマッチを認識できることを示した（図１０、１１）。これらのタンパクはアミノ酸配列中に大腸菌MutMタンパクとの明確なホモロジーはない（van der Kemp *et al.* 1996, Nash *et al.* 1996）。この事実は、これらのタンパクのC-Cのミスマッチの認識と結合がアーキファクトではなく、必須の機能であることを示唆している。さらに、*S. cerevisiae* の粗抽出液中にはOGG1とは異なると思われる、C-Cミスマッチ結合タンパクの存在が確認された（図１２）。このタンパクはMutM/OGG1と同じように、シトシンを含むミスマッチに結合することができる。さらに、アデニン同士のミスマッチにも結合できることも分かった。

## 大腸菌の GC-CG の変異を特異的にもたらすミューテーター遺伝子の同定

現在までに大腸菌で見つかっている様々なミューテーター株には、CC103株の自然突然変異頻度だけを上昇させるようなものは見つかっていない。そこで、大腸菌CC103株（Cupples *et al.* 1989）（図１３）に直接トランスポゾンを導入することで、ランダムに遺伝子を破壊した菌株を作成し、その中からGC-CGトランスバー

ジョンの自然突然変異頻度が高くなるような菌株（ミューテーター株）を探索した。約４万個のトランスポゾン挿入コロニーを調べた結果、５つの候補が得られた（図１４）。これらをさらに調べた結果、３つのグループに分けることができた（表３）。グループ１はGC-CGトランスバージョンが起こったことを示すCC103株でのパピーレの発生頻度と、6種類の塩基置換型突然変異をすべて含んでいると考えられるリファンピシン耐性菌の発生頻度がほぼ同じである。これに対して、グループの２と３はリファンピシン耐性菌の発生頻度がCC103株でのパピーレの発生頻度を大きく上回っているため、GC-CG以外の突然変異が多く起こっていることを示している。このことから、今回求めている、GC-CGの突然変異特異的なミューテーター株はグループ１のGC39株だけであると判断した。

次に、SingerのP1マッピングキットの大腸菌（Singer *et al.* 1989）を使って、GC39株でトランスポゾンによって破壊されいる、ミューテーター遺伝子 *mut39* の位置をマッピングした。*mut39*遺伝子と同時に遺伝子が移る割合は67.9分の遺伝子で23 %、66.0分の遺伝子で19 %、66.9分の遺伝子で90 %であった。このことから、*mut39*遺伝子は大腸菌ゲノム上の67分付近にあることが分かった。67分付近にはDNA修復酵素として知られている、*mutY*遺伝子もマップされている（Nghiem *et al.* 1988）。そこで、miniTn*10* の末端領域の配列を利用して、トランスポゾンが挿入されている部分の周囲のDNA配列をシークエンスしてみた。得られたシークエンスで遺伝学研究所の大腸菌のORFデータベースを検索したところ、 *mutY*遺伝子中に存在していた。このことは *mut39*変異は *mutY*遺伝子のアリルであることを示している。また、大腸菌CC104*mut39*株でCC104 *mutY*株と同じようにGC-TAの突然変異頻度が上昇することもその事実と矛盾しない（表４）。その上、CC103*mutY*株でも $Lac^+$ の復帰突然変異の頻度が上昇することも分かった（表４）。

大腸菌の８－オキソグアニンに対する防御機構である GO システムとの関連

プリン塩基は環状部分の酸化によって様々な化学的修飾を受ける。それらの中で、高い突然変異源性を持つ8－オキソグアニンは生成量が多い（Bridges eta al. 1996, Cheng *et al.* 1992）。様々なDNAポリメラーゼが8－オキソグアニンの向かいに高い確率でアデニンを挿入することができる（Wood *et al.* 1990, Shibutani *et al.* 1991, Tchou *et al.* 1993）。そのため、鋳型鎖中に8－オキソグアニンが存在するとGC-TAの突然変異が生じうる（Tajiri *et al.* 1995, Moriya *et al.* 1993, Cheng *et al.* 1992, Grollman *et al.* 1993）。大腸菌での研究では、MutM、MutY、MutTの3つの修復酵素が共同して8－オキソグアニンから生じる突然変異を抑制していることがわかっている。MutMタンパクは開環したり、開環していないが酸化したプリン塩基、FaPy塩基や8－オキソグアニンをDNAから除去するグリコシラーゼである（Demple *et al.* 1994, Tchou *et al.* 1991, Michaels *et al.* 1991, Rabow *et al.* 1997）。MutYタンパクは8－オキソグアニンと対合したアデニンを除去するグリコシラーゼである（Demple *et al.* 1994, Michaels *et al.* 1990a, Michaels *et al.* 1990b, Michaels *et al.* 1992）。このミスマッチは、鋳型鎖に8－オキソグアニンが残っているときにDNA複製が起こると生じる。また、G-Aミスマッチからもアデニンを効率よく除去する（Aboul-ela *et al.* 1985）。MutTタンパクは、DNA複製時に取り込まれるのを防ぐため、8－オキソデオキシグアノシン3リン酸を8－オキソデオキシグアノシン1リン酸に加水分解する酵素である（Maki *et al.* 1992）。*mutMmutY*の2重変異株ではGC-TAの突然変異頻度が上昇する（Nghiem *et al.* 1988, Tajiri et al 1995, Moriya *et al.* 1993）。

今回の実験で*mutY*変異株でGC-CGの突然変異頻度が上昇することが示された（表4）。この結果から2つのモデルを考えてみた。第1に、8－オキソグアニンのような損傷塩基がグルコース飢餓状態の細胞に生じ、*mutY*欠損時に多くのGC-CGの突然変異を発生させるというものである。定常期の細胞は増殖期の細胞よりも約3倍の8－オキソグアニンを産生している（Bridges *et al.* 1996）。すなわち、

８－オキソグアニンは大腸菌の定常期の突然変異の重要な構成要素であるのかもしれない。DNA 複製時に８－オキソグアニンの向かいに dAMP だけでなく dGMP も挿入されうることが Braun ら（Braun *et al.* 1997）によって示唆されている。８－オキソグアニンとグアニンの塩基対が修復されなければ、GC-CG の突然変異の原因となる。第２に、大腸菌において、MutY タンパクが DNA 中の８－オキソグアニンとグアニンの塩基対を認識、修復する事でき、GC-CG の突然変異の発生を押さえているというものである。事実、精製した MutY タンパクは、８－オキソグアニンとグアニンの塩基対に、８－オキソグアニンとアデニンの塩基対と同じくらい強く結合することができ（図１５）、またグアニン-DNA グリコシラーゼ活性も持っている（図１６）。しかし、その除去の反応速度は ８－オキソグアニンとアデニンの塩基対の場合よりも遅くなっている（図１７、１８）。

*mutM* 変異を導入した CC103 株や CC103 *mutY* 株では Lac$^+$ の復帰突然変異の頻度は上昇しない（表４）。MutM タンパクは *in vitro* で８－オキソグアニンとグアニンの塩基対から８－オキソグアニンを除去することができることが示されている（Michaels *et al.* 1992a,Tchou *et al.* 1991）。しかし、MutM タンパクが８－オキソグアニンとグアニンの塩基対から８－オキソグアニンを除去すると、GC-CG の突然変異の原因となるかもしれない。なぜなら、DNA 合成時に８－オキソグアニンに向かいに dGMP が挿入された箇所では、修復合成の時に DNA 中に残っているグアニンの向かいに dCMP が挿入されるからである。そのため、大腸菌において MutY タンパクが８－オキソグアニンとグアニンの塩基対からグアニンを除去することで、同じ塩基対から MutM タンパクによる８－オキソグアニンの除去を抑制することは合理的である。このような MutM タンパクの活性の制御に MutY タンパクが関係している可能性を Bridges ら（Bridges *et al.* 1996）が示唆している。また、Braun ら（Braun *et al.* 1997）はヌクレオチド除去修復系が酸化的損傷を受けたグアニンと正常なグアニンの塩基対から損傷したグアニンを除去することができ、GC-CG の突然

変異の原因となっていることを示唆している。

GC-CG 変異の発現の特異性

Nghiem らはパピーレを作らせる方法を用いて（図１３、１４）、*mutY*変異によって GC-TA の自然突然変異頻度は上昇するがその他の突然変異の頻度は上昇しないことを報告している（Nghiem *et al.* 1988）。彼らの実験では、大腸菌をグルコース最小培地で３日培養してパピーレを観察している。本研究では大腸菌を１０日培養している。このような条件下では、多少の代謝回転とゆっくりとした細胞数の増加が観察されている（Bridges *et al.* 1996）。GC-CG の突然変異が遅れて増えてくる現象は Mackay らも報告している（Mackay *et al.* 1994）。８－オキソグアニンなどの損傷は常に生成し続けており、鋳型 DNA 上の損傷に dCMP だけでなく dAMP や dGMP が間違って挿入されていると考えられる。しかし、MutY タンパクによる８－オキソグアニンとグアニンの塩基対の修復は８－オキソグアニンとアデニンの塩基対の修復よりも効率が低い（図１７、１８）。そのため、GC-TA の突然変異よりも GC-CG の突然変異の方が表現型として現れるのに時間がかかると考えられる。

他の生物の MutY のホモログについて

大腸菌 MutY 遺伝子のヒトホモログ（ *hMYH*）がクローニングされて塩基配列が決定された（Slupska *et al.* 1996）。*hMYH* 遺伝子は５３５アミノ酸のタンパクをコードしている。この大きさは抗大腸菌 MutY 抗体による交差反応で HeLa 細胞から得られるポリペプチドの大きさとよく一致している（McGoldrick *et al.* 1995）。ヒト *hMYH* 遺伝子は大腸菌 MutY タンパクと 41 %の相同性があり、DNA の酸化的損傷の修復と酸化的損傷による突然変異の抑制に対する重要な機能を示唆している。

MutM/OGG1 や FabA/FAS で示されたような、機能の相同性の有無は今後の検討課題である。

# 謝辞

本研究を行うにあたりご協力いただきました以下の皆様に深く感謝いたします。様々な大腸菌の変異株やプラスミドを分与していただいた、国立癌研究所の横田淳博士、放射線医学総合研究所の味村正博博士、東北大学理学部の山本和生教授、ミシガン大学のB. Weiss博士、J. H. Miller博士、国立遺伝学研究所の西村昭子助教授、ハーバード大学のM. M. Wu博士、カロリンスカ大学のA. Holmgren博士、エール大学のB. Bachmann博士に深くお礼を申し上げます。

また、京都大学放射線生物学研究センターの藤堂剛助教授にはゲルシフトアッセイなどについてご助言をいただき、深く感謝いたします。

指導教官として、また共同研究者として様々なご指導とご助言をいただき、さらに本稿の精読と批評をしていただきました米井脩治教授、張秋梅助手に心よりお礼を申し上げます。最後に、実験の指導をしていただいたり、研究内容について議論していただいた動物学教室放射線生物学講座の皆様、また当研究室の卒業生の皆様に深くお礼を申し上げます。

# 引用文献

Aboul-ela, F., Koh, D., Tinoco, I., Jr and Martin, F. H. (1985). Base-base mismatches. Thermodynamics of double helix formation for dCA3XA3G + dCT3YT3G (X, Y = A,C,G,T). *Nucleic Acids Res.* 13 , 4811-4824.

Akasaka, S.and Yamamoto, K. (1994). Hydrogen peroxide induces G:C to T:A and G:C to C:G transversions in the supF gene of *Escherichia coli. Mol. Gen. Genet.* 243 , 500-505.

Au, K.G., Clark S., Miller, J.H. and Modrich, P. (1989). *Escherichia coli mutY* gene encodes an adenine glycosylase active on G-A mispairs. *Proc. Natl. Acad. Sci. USA.* 86 , 8877-8881.

Babic, I., Andrew, S.E. and Jirik F.R. (1996). MutS interaction with mismatch and alkylated base containing DNA molecules detected by optical biosensor. *Mutat. Res.* 372 , 87-96.

Bjoras, M., Luna, L., Johnsen, B., Hoff, E., Haug, T., Rognes, T. and Seeberg, E. (1997). Opposite base-dependent reactions of a human base excision repair enzyme on DNA containing 7,8-dihydro-8-oxoguanine and abasic sites. *EMBO J.* 16 , 6314-6322.

Boiteux, S., O'Connor, T.R. and Laval, J. (1987). Formamidopyrimidine-DNA glycosylase of *Escherichia coli*: cloning and sequencing of the fpg structural gene and overproduction of the protein. *EMBO J.* 6, 3177-3183.

Bridges, B.A., Sekiguchi, M. and Tajiri., T. (1996). Effect of *mutY* and *mutM/fpg-1* mutations on starvation-associated mutation in *Escherichia coli* :implications for the role of

7,8-dihydro-8-oxoguanine. *Mol. Gen. Genet.* 251, 352-357.

Buchko, G.W., Wagner, J.R., Cadet, J., Raoul, S. and Weinfeld, M. (1995). Methylene blue-mediated photooxidation of 7,8-dihydro-8-oxo-2'-deoxyguanosine. *Biochim. Biophys. Acta.* 1263, 17-24.

Carlioz, A. and Touati, D. (1986). Isolation of superoxide dismutase mutants in *Escherichia coli*: is superoxide dismutase necessary for aerobic life? *EMBO J.* 5, 623-630.

Cabrera, M., Nghiem, Y. and Miller, J.H. (1988). *mutM*, a second mutator locus in *Escherichia coli* that generates G.C----T.A transversions. *J. Bacteriol.* 170, 5405-5407.

Cheng, K.C., Cahill, D.S., Kasai, H., Nishimura, S. and Loeb L.A. (1992). 8-Hydroxyguanine, an abundant form of oxidative DNA damage, causes G----T and A----C substitutions. *J. Biol. Chem.* 267, 166-172.

Cronan, J.E. Jr, Li, W.B., Coleman, R., Narasimhan, M., de Mendoza, D. and Schwab, J.M. (1988). Derived amino acid sequence and identification of active site residues of *Escherichia coli* $\beta$-hydroxydecanoyl thioester dehydrase. *J. Biol. Chem.* 263, 4641-4646.

Cupples, C.G. and Miller, J.H. (1989). A set of *lacZ* mutations in *Escherichia coli* that allow rapid detection of each of the six base substitutions. *Proc. Natl. Acad. Sci. USA.* 86, 5345-5349.

Dherin, C., Radicella, J.P., Dizdaroglu, M. and Boiteux, S. (1999). Excision of oxidatively

damaged DNA bases by the human alpha-hOgg1 protein and the polymorphic alpha-hOgg1 (Ser326Cys) protein which is frequently found in human populations. *Nucleic Acids Res.* 27, 4001-4007.

Demple, B. and Harrison, L. (1994). Repair of oxidative damage to DNA: enzymology and biology. *Annu. Rev. Biochem.* 63, 915-948.

Drake, J.W. (1991). Spontaneous mutation. *Annu. Rev. Genet.* 25, 125-146.

Echols, H. and Goodman, M.F. (1991). Fidelity mechanisms in DNA replication. *Annu. Rev. Biochem.* 60, 477-511.

Modrich, P. (1991). Mechanisms and biological effects of mismatch repair. *Annu. Rev. Genet.* 25, 229-253.

Eisen, J.A and Hanawalt, P.C. (1999). A phylogenomic study of DNA repair genes, proteins, and processes. *Mutat. Res.* 435, 171-213.

Fleck, O., Schar, P. and Kohli, J. (1994). Identification of two mismatch-binding activities in protein extracts of *Schizosaccharomyces pombe. Nucleic Acids Res.* 22, 5289-5295.

Fleck, O., Kunz, C., Rudolph, C. and Kohli, J. (1998). The high mobility group domain protein Cmb1 of *Schizosaccharomyces pombe* binds to cytosines in base mismatches and opposite chemically altered guanines. *J. Biol. Chem.* 273, 30398-30405.

Fleck, O., Lehmann, E., Schar, P. and Kohli, J. (1999). Involvement of nucleotide-excision repair in *msh2 pms1*-independent mismatch repair. *Nat. Genet.* 21, 314-317.

Fox, M.S., Radicella, J.P. and Yamamoto, K. (1994). Some features of base pair mismatch repair and its role in the formation of genetic recombinants. *Experientia.* 50. 253-260.

Girard, P.M., Guibourt, N. and Boiteux, S. (1997). The Ogg1 protein of *Saccharomyces cerevisiae*: a 7,8-dihydro-8-oxoguanine DNA glycosylase/AP lyase whose lysine 241 is a critical residue for catalytic activity. *Nucleic Acids Res.* 25, 3204-3211.

Grollman, A.P. and Moriya, M. (1993). Mutagenesis by 8-oxoguanine: an enemy within. *Trends. Genet.* 9, 246-249.

Halling, S.M., Simons, R.W., Way, J.C., Walsh, R.B. and Kleckner, N. (1982). DNA sequence organization of IS10-right of Tn10 and comparison with IS10-left. *Proc. Natl. Acad. Sci. USA.* 79, 2608-2612.

Hatahet, Z., Kow, Y.W., Purmal, A.A., Cunningham, R.P. and Wallace, S.S. (1994). New substrates for old enzymes. 5-Hydroxy-2'-deoxycytidine and 5-hydroxy-2'-deoxyuridine are substrates for *Escherichia coli* endonuclease III and formamidopyrimidine DNA N-glycosylase, while 5-hydroxy-2'-deoxyuridine is a substrate for uracil DNA N-glycosylase. *J. Biol. Chem.* 269, 18814-18820.

Ishida, T., Hippo, Y., Nakahori, Y., Matsushita, I., Kodama, T., Nishimura, S. and Aburatani, H. (1999). Structure and chromosome location of human *OGG1*. *Cytogenet. Cell. Genet.* 85,

232-236.

Kaslan, E. and Heyer, W.D. (1994). *Schizosaccharomyces pombe* fatty acid synthase mediates DNA strand exchange *in vitro*. *J. Biol. Chem.* 269, 14103-14110.

Jiricny, J., Hughes, M., Corman, N. and Rudkin, B. B. (1988). A human 200-kDa protein binds selectively to DNA fragments containing G-T mismatches. *Proc. Natl. Acad. Sci. USA.* 85, 8860-8864.

van der Kemp, P.A., Thomas, D., Barbey, R., de Oliveira, R. and Boiteux, S. (1996). Cloning and expression in *Escherichia coli* of the *OGG1* gene of *Saccharomyces cerevisiae*, which codes for a DNA glycosylase that excises 7,8-dihydro-8-oxoguanine and 2,6-diamino-4-hydroxy-5-N-methylformamidopyrimidine. *Proc. Natl. Acad. Sci. USA.* 93, 5197-5202.

Kleckner, N., Barker, D.F., Ross, D.G. and Botstein, D. (1978). Properties of the translocatable tetracycline-resistance element Tn *10* in *Escherichia coli* and bacteriophage lambda. *Genetics.* 90, 427-461.

Kramer, B., Kramer, W. and Fritz, H.J. (1984). Different base/base mismatches are corrected with different efficiencies by the methyl-directed DNA mismatch-repair system of *E. coli*. *Cell.* 38, 879-887.

Mackay, W.J., Han, S. and Samson, L.D. (1994). DNA alkylation repair limits spontaneous base substitution mutations in *Escherichia coli*. *J. Bacteriol.* 176, 3224-3230.

Magnuson, K., Jackowski, S., Rock, C.O. and Cronan, J.E. Jr. (1993). Regulation of fatty acid biosynthesis in *Escherichia coli. Microbiol. Rev.* 57, 522-542.

Maki, H. and Sekiguchi, M. (1992). MutT protein specifically hydrolyses a potent mutagenic substrate for DNA synthesis. *Nature.* 355, 273-275.

McBride, T.J., Preston, B.D. and Loeb, L.A. (1991). Mutagenic spectrum resulting from DNA damage by oxygen radicals. *Biochemistry.* 30, 207-213.

McGoldrick, J.P., Yeh, Y.C., Solomon, M., Essigmann, J.M. and Lu, A.L. (1995). Characterization of a mammalian homolog of the *Escherichia coli* MutY mismatch repair protein. *Mol. Cell. Biol.* 15, 989-996.

Michaels, M.L., Pham, L., Nghiem, Y., Cruz, C. and Miller, J.H. (1990). MutY, an adenine glycosylase active on G-A mispairs, has homology to endonuclease III. *Nucleic Acids Res.* 18, 3841-3845.

Michaels, M.L., Cruz, C. and Miller, J.H. (1990). *mutA* and *mutC*: two mutator loci in *Escherichia coli* that stimulate transversions. *Proc. Natl. Acad. Sci. USA.* 87, 9211-9215.

Michaels, M.L., Pham, L., Cruz, C. and Miller, J.H. (1991). MutM, a protein that prevents G.C----T.A transversions, is formamidopyrimidine-DNA glycosylase. *Nucleic Acids Res.* 19, 3629-3632.

Michaels, M.L., Cruz, C., Grollman, A.P. and Miller, J.H. (1992). Evidence that MutY and

MutM combine to prevent mutations by an oxidatively damaged form of guanine in DNA. *Proc. Natl. Acad. Sci. USA.* 89, 7022-7025.

Michaels, M.L. and Miller, J.H. (1992). The GO system protects organisms from the mutagenic effect of the spontaneous lesion 8-hydroxyguanine (7,8-dihydro-8-oxoguanine). *J. Bacteriol.* 174, 6321-6325.

Miller, J.H. (1996). Spontaneous mutators in bacteria: insights into pathways of mutagenesis and repair. *Annu. Rev. Microbiol.* 50, 625-643.

Miller, J.H. (1998). Mutators in *Escherichia coli. Mutat. Res.* 409, 99-106.

Miret, J. J., Milla, M. G. and Lahue, R. S. (1993). Characterization of mismatch-binding activity in yeast extracts. *J. Biol. Chem.* 268, 3507-3513.

Modrich, P. (1991). Mechanisms and biological effects of mismatch repair. *Annu. Rev. Genet.* 25, 229-253.

Modrich, P. and Lahue, R. (1996). Mismatch repair in replication fidelity, genetic recombination, and cancer biology. *Annu. Rev. Biochem.* 65, 101-133.

Mol, C.D., Parikh, S.S., Putnam, C.D., Lo, T.P. and Tainer, J.A. (1999). DNA repair mechanisms for the recognition and removal of damaged DNA bases. *Annu. Rev. Biophys. Biomol. Struct.* 28, 101-128.

Moriya, M. and Grollman, A.P. (1993). Mutations in the *mutY* gene of *Escherichia coli* enhance the frequency of targeted G:C-->T:A transversions induced by a single 8-oxoguanine residue in single-stranded DNA. *Mol. Gen. Genet.* 239, 72-76.

Nagashima, M., Sasaki, A., Morishita, K., Takenoshita, S., Nagamachi, Y., Kasai, H. and Yokota, J. (1997). Presence of human cellular protein(s) that specifically binds and cleaves 8-hydroxyguanine containing DNA. *Mutat. Res.* 383, 49-59.

Nash, H.M., Bruner, S.D., Scharer, O.D., Kawate, T., Addona, T.A., Spooner, E., Lane, W.S. and Verdine, G.L. (1996). Cloning of a yeast 8-oxoguanine DNA glycosylase reveals the existence of a base-excision DNA-repair protein superfamily. *Curr. Biol.* 6, 968-980.

Nghiem, Y., Cabrera, M., Cupples, C.G. and Miller, J.H. (1988). The *mutY* gene: a mutator locus in *Escherichia coli* that generates G.C----T.A transversions. *Proc. Natl. Acad. Sci. USA.* 85, 2709-2713.

Nishimura, A., Akiyama, K., Kohara, Y. and Horiuchi, K. (1992). Correlation of a subset of the pLC plasmids to the physical map of *Escherichia coli* K-12. *Microbiol. Rev.* 56, 137-151.

Purmal, A. A., Kow, Y. W. and Wallace, S. S. (1994). Major oxidative products of cytosine, 5-hydroxycytosine and 5-hydroxyuracil, exhibit sequence context-dependent mispairing *in vitro. Nucleic Acids Res.* 22, 72-78.

Purmal, A. A., Kow, Y. W. and Wallace, S. S. (1994). 5-Hydroxypyrimidine

deoxynucleoside triphosphates are more efficiently incorporated into DNA by exonuclease-free Klenow fragment than 8-oxopurine deoxynucleoside riphosphates. *Nucleic Acids Res.* 22, 3930-3935.

Rabow, L.E. and Kow, Y.W. (1997). Mechanism of action of base release by *Escherichia coli* Fpg protein: role of lysine 155 in catalysis. *Biochemistry.* 36, 5084-5096.

Radicella, J.P., Clark, E.A. and Fox, M.S. (1988). Some mismatch repair activities in *Escherichia coli. Proc. Natl. Acad. Sci. USA.* 85, 9674-9678.

Saito, K., Hamajima, A., Ohkuma, M., Murakoshi, I., Ohmori, S., Kawaguchi, A., Teeri, T.H. and Cronan, J.E. Jr. (1995). Expression of the *Escherichia coli fabA* gene encoding $\beta$-hydroxydecanoyl thioester dehydrase and transport to chloroplasts in transgenic tobacco. *Transgenic. Res.* 4, 60-69.

Sandigursky, M., Yacoub, A., Kelley, M.R., Xu, Y., Franklin, W.A. and Deutsch, W.A. (1997). The yeast 8-oxoguanine DNA glycosylase (Ogg1) contains a DNA deoxyribophosphodiesterase (dRpase) activity. *Nucleic Acids Res.* 25, 4557-4561.

Sargentini, N.J. and Smith, K.C. (1994). DNA sequence analysis of $\gamma$-radiation (anoxic)-induced and spontaneous *lacI^d* mutations in *Escherichia coli* K-12. *Mutat. Res.* 309, 147-163.

Schaaper, R.M. and Dunn, R.L. (1987). Spectra of spontaneous mutations in *Escherichia coli* strains defective in mismatch correction: the nature of *in vivo* DNA replication errors. *Proc.*

*Natl. Acad. Sci. USA.* 84, 6220-6224.

Schaaper, R.M. (1993). Base selection, proofreading, and mismatch repair during DNA replication in *Escherichia coli. J. Biol. Chem.* 268, 23762-23765.

Schar, P., Munz, P. and Kohli, J. (1993). Meiotic mismatch repair quantified on the basis of segregation patterns in *Schizosaccharomyces pombe. Genetics.* 133, 815-824.

Schar, P. and Kohli, J. (1993). Marker effects of G to C transversions on intragenic recombination and mismatch repair in *Schizosaccharomyces pombe*. *Genetics.* 133, 825-835.

Shibutani, S., Takeshita, M. and Grollman, A.P. (1991). Insertion of specific bases during DNA synthesis past the oxidation-damaged base 8-oxodG. *Nature.* 349, 431-434.

Shinmura, K., Kasai, H., Sasaki, A., Sugimura, H. and Yokota, J. (1997). 8-hydroxyguanine (7,8-dihydro-8-oxoguanine) DNA glycosylase and AP lyase activities of hOGG1 protein and their substrate specificity. *Mutat. Res.* 385, 75-82.

Silbert, D.F. and Vagelos, P.R. (1967). Fatty acid mutant of *E. coli* lacking a $\beta$-hydroxydecanoyl thioester dehydrase. *Proc. Natl. Acad. Sci. USA.* 58, 1579-1586.

Singer, M., Baker, T.A., Schnitzler, G., Deischel, S.M., Goel, M., Dove, W., Jaacks, K.J., Grossman, A.D., Erickson, J.W. and Gross, C.A. (1989). A collection of strains containing genetically linked alternating antibiotic resistance elements for genetic mapping of

*Escherichia coli. Microbiol. Rev.* 53, 1-24.

Slupska, M.M., Baikalov, C., Luther, W.M., Chiang, J.H., Wei, Y.F. and Miller, J.H. (1996). Cloning and sequencing a human homolog (*hMYH*) of the *Escherichia coli mutY* gene whose function is required for the repair of oxidative DNA damage. *J. Bacteriol.* 178, 3885-3892.

Smith, K.C. (1992). Spontaneous mutagenesis: experimental, genetic and other factors. *Mutat. Res.* 277, 139-162.

Su, S.S., Lahue, R.S., Au, K.G. and Modrich, P. (1988). Mispair specificity of methyl-directed DNA mismatch correction *in vitro. J. Biol. Chem.* 263, 6829-6835.

Tajiri, T., Maki, H. and Sekiguchi, M. (1995). Functional cooperation of MutT, MutM and MutY proteins in preventing mutations caused by spontaneous oxidation of guanine nucleotide in *Escherichia coli. Mutat. Res.* 336, 257-267.

Tchou, J., Kasai, H., Shibutani, S., Chung, M.H., Laval, J., Grollman, A.P. and Nishimura, S. (1991). 8-oxoguanine (8-hydroxyguanine) DNA glycosylase and its substrate specificity. *Proc. Natl. Acad. Sci. USA.* 88, 4690-4694.

Tchou, J. and Grollman, A.P. (1993). Repair of DNA containing the oxidatively-damaged base, 8-oxoguanine. *Mutat. Res.* 299, 277-287.

Tchou, J., Bodepudi, V., Shibutani, S., Antoshechkin, I., Miller, J., Grollman, A.P. and

Johnson, F. (1994). Substrate specificity of Fpg protein. Recognition and cleavage of oxidatively damaged DNA. *J. Biol. Chem.* 269, 15318-15324.

Tsai-Wu, J.J., Radicella, J.P. and Lu, A.L. (1991). Nucleotide sequence of the *Escherichia coli micA* gene required for A/G-specific mismatch repair:identity of *micA* and *mutY*. *J. Bacteriol.* 173, 1902-1910.

Wagner, R., Debbie, P. and Radman, M. (1995). Mutation detection using immobilized mismatch binding protein (MutS). *Nucleic Acids Res.* 23, 3944-3948.

Wood, M.L., Dizdaroglu, M., Gajewski, E. and Essigmann, J.M. (1990). Mechanistic studies of ionizing radiation and oxidative mutagenesis: genetic effects of a single 8-hydroxyguanine (7-hydro-8-oxoguanine) residue inserted at a unique site in a viral genome. *Biochemistry.* 29, 7024-7032.

Worth, L. Jr, Clark, S., Radman, M. and Modrich, P. (1994). Mismatch repair proteins MutS and MutL inhibit RecA-catalyzed strand transfer between diverged DNAs. *Proc. Natl. Acad. Sci. USA.* 91, 3238-3241.

Zhang, Q.-M., Sugiyama, H., Miyabe, I., Matsuda, S., Saito, I. and Yonei, S. (1997). Replication of DNA templates containing 5-formyluracil, a major oxidative lesion of thymine in DNA. *Nucleic Acids Res.* 25, 3969-3973.

--------------------------------------------------------------------

Miller, J. H. (1992). A Short Course in Bacterial Genetics. Cold Spring Harbor Laboratory

Press, NY.

Berlyn, M. K. B., Low, K. B. and Rudd, K. E. (1996). In Neidhardt, F. C. (ed.). *Escherichia coli* and Salmonella, Vol. 2. ASM Press, WA, pp. 1715-1902.

Singer, M. and Low, K. B. (1996). In Neidhardt, F. C. (ed.) *Escherichia coli* and *Salmonella*, Vol. 2. ASM Press, WA, pp. 1812-1815.

Wallace, S. S. (1997). Oxidative Stress and the Molecular Biology of Antioxidant Defenses. (ed. J. G. Scandalios) Cold Spring Harbor Laboratory Press, NY, pp. 49-90.

Fritsch, E.F., Maniatis, T. and Sambrook, J. (1989). Molecular Cloning : A Laboratory Manual. Cold Spring Harbor Laboratory Press, NY.

Friedberg, E. C., Walker, G. C. and Siede, W. (1995). DNA repair and mutagenesis. ASM Press, WA.

# 図と表

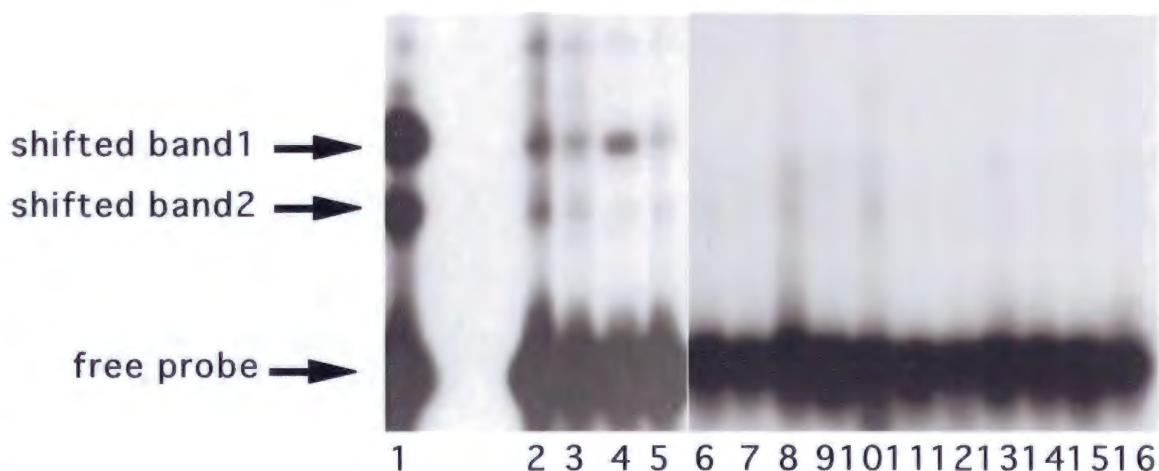



図１．大腸菌野生株の粗抽出液中に存在するミスマッチ結合活性の確認。

$^{32}$P でラベルした様々なミスマッチを含む２本鎖のオリゴヌクレオチド 23.5 fmolを大腸菌野生株の粗抽出液に加え、4℃で 30 分反応させた。反応液を 12 %の非変性ポリアクリルアミドゲル、1 X TBE バッファー上で 100 V で電気泳動した。ゲルを乾燥させ、-80℃で X 線フィルムに感光させた。各レーンに用いられているプローブは、レーン１，C/C、レーン２，T/C、レーン３，A/C、レーン４，C/T、レーン５，C/A、レーン６，C/G、レーン７，A/A、レーン８，A/G、レーン９，A/T、レーン１０，G=C、レーン１１，G/A、レーン１２，G/G、レーン１３，G/T、レーン１４，T=A、レーン１５，T/G、レーン１６，T/T。ラベルした DNA とその相補鎖の組み合わせは以下の通り。レーン１、４～６は図２の１の DNA をラベルし、５～８の DNA とアニールさせた。レーン２、１４～１６は４の DNA をラベルし、５～８の DNA とアニールさせた。レーン３、７～９は２の DNA をラベルし、５～８の DNA とアニールさせた。レーン１０～１３は３の DNA をラベルし、５～８の DNA とアニールさせた。

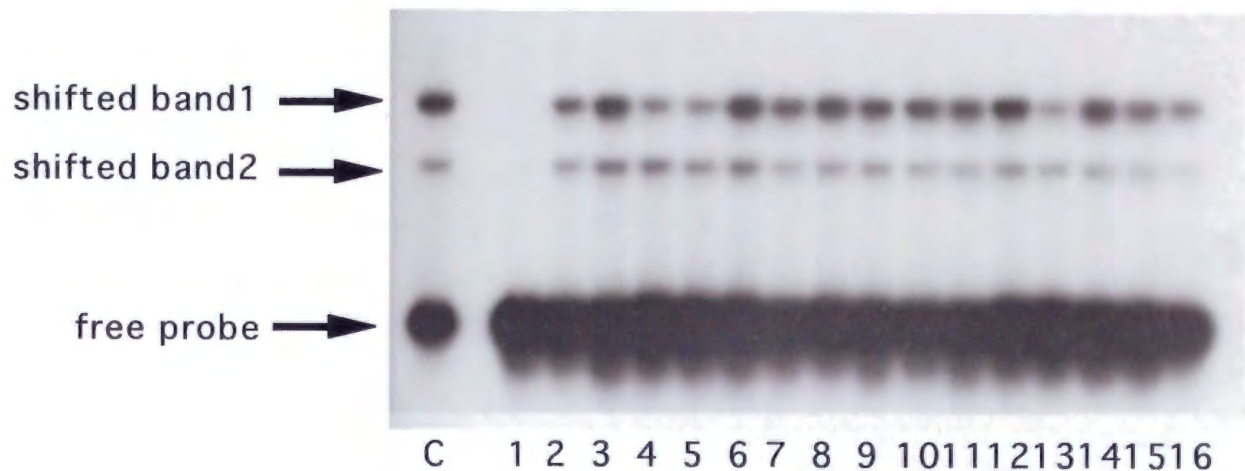



図２．C/Cミスマッチに対する特異性をコンペティションアッセイで確認。$^{32}$P でラベルしたC/Cミスマッチを含むプローブを含む反応液に、その 50 倍量のラベルしていないプローブをコンペティターとして加え、電気泳動した。泳動の条件は図３と同じである。各レーンに加えられたコンペティターは、レーンC,コンペティターなし、レーン１，C/C、レーン２，C/A、レーン３，C=G、レーン４，C/T、レーン５，A/C、レーン６，A/A、レーン７，A/G、レーン８，A=T、レーン９，G=C、レーン１０，G/A、レーン１１，G/G、レーン１２，G/T、レーン１３，T/C、レーン１４，T=A、レーン１５，T/G、レーン１６，T/T。各組み合わせのコンペティターの作り方は図１を参照。

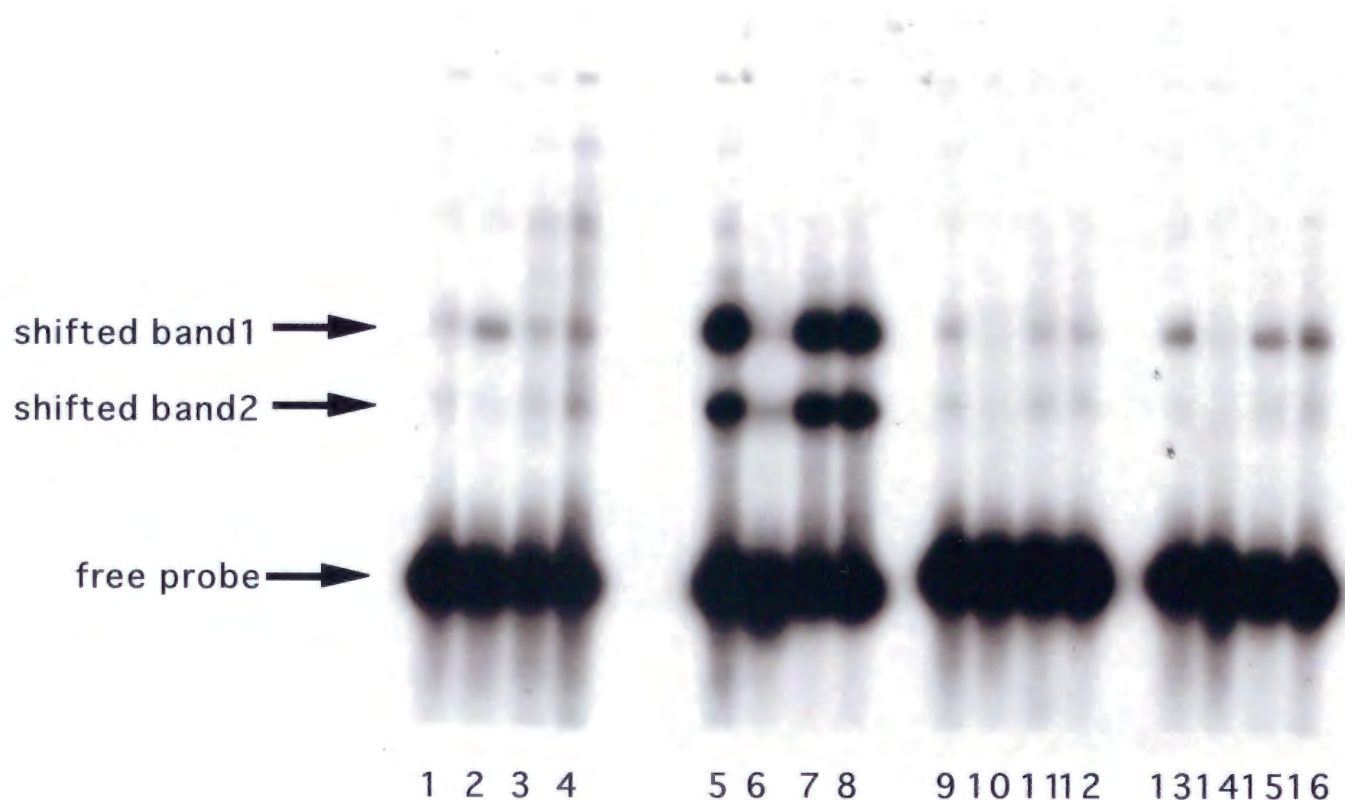



図３．シトシンを含むミスマッチ間の親和性の比較

C/C以外のシトシンを含むミスマッチとC/Cのミスマッチとの親和性の違いを比較した。各レーンのプローブは、レーン１，C/A、レーン２，C/T、レーン３，A/C、レーン４，T/C、レーン５～８，C/C、レーン９～１２，C/A、レーン１３～１６，C/T。また、レーン６、１０、１４にはC/C、レーン７、１１、１５にはC/A、レーン８、１２、１６にはC/Tのコンペティターがプローブの50倍加えられている。反応や電気泳動の条件は図１と同じ。

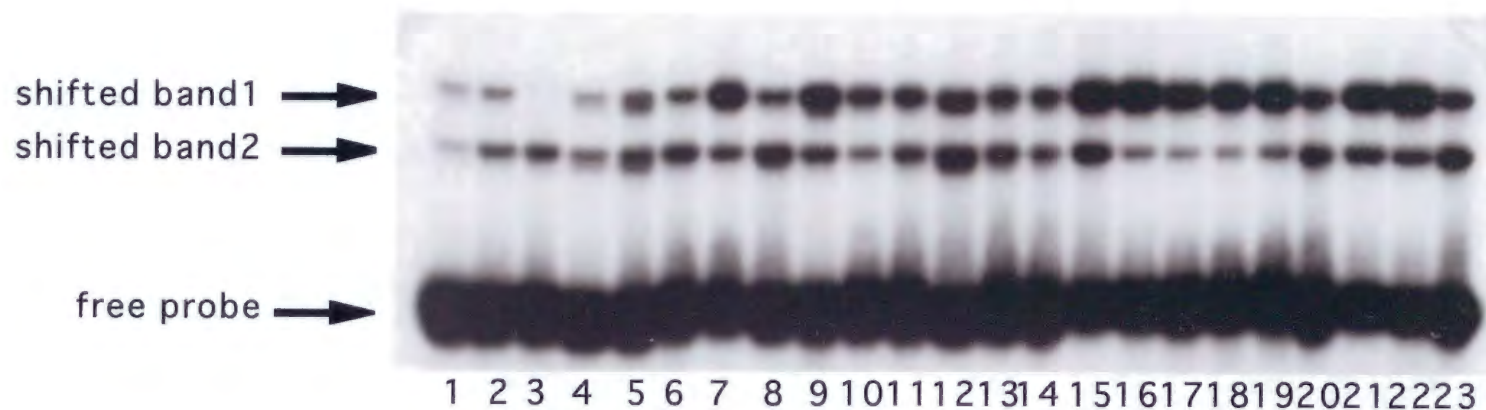



図４．様々な変異株の粗抽出液中のＣ／Ｃミスマッチ結合活性の確認。

Ｃ／Ｃのミスマッチに結合する2つのタンパクを同定するために、まず様々な大腸菌の変異株の粗抽出液を用いたゲルシフトアッセイを行い、シフトしたバンドが無くなるような変異株を探した。プローブはすべてＣ／Ｃのミスマッチのものを用いた。各レーンで用いた変異株で欠損している遺伝子は、レーン１，*mutD*、レーン２，*mutL*、レーン３，*mutM*、レーン４，*mutT*、レーン５，*mutY*、レーン６，*ung*、レーン７，*vsr*、レーン８，*uvrD*、レーン９，*ada ogt*、レーン１０，*xth nfo*、レーン１１，*alkA*、レーン１２，*hupA hupB*、レーン１３，*dps*、レーン１４，*mfd*、レーン１５，*nth nei*、レーン１６，*recA*、レーン１７，*mutH*、レーン１８，*mutS*、レーン１９，*recB recC sbcB*、レーン２０，*topA*、レーン２１，*umuC umuD*、レーン２２，*uvrA*、レーン２３，*topB*。電気泳動などの条件は図１と同じ。

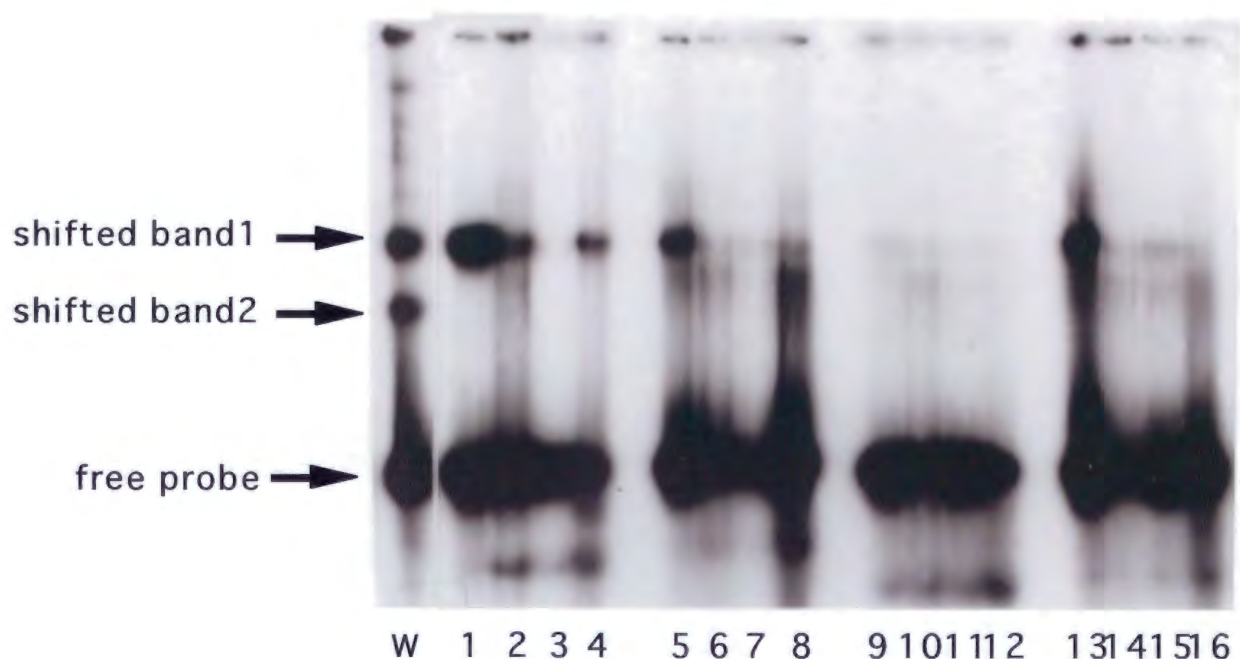



図５．精製した大腸菌 MutM タンパクのミスマッチ認識の特異性の確認
大腸菌 BL21 株に *GST-MutM* 融合遺伝子を持つプラスミドを導入し、IPTG 処理で過剰発現させた。この融合タンパクを Glutathione-Sepharose 4B カラムを使って精製した。精製した GST-MutM タンパクをトロンビンプロテアーゼで分解して GST-tag をはずした。この MutM タンパクを用いて図３と同じ実験を行った。各レーンに用いられているプローブは、レーンW，Ｃ/Ｃのプローブに大腸菌野生株の粗抽出液を加えたもの、レーン１，Ｃ/Ｃ、レーン２，Ｃ/Ａ、レーン３，Ｃ=Ｇ、レーン４，Ｃ/Ｔ、レーン５，Ａ/Ｃ、レーン６，Ａ/Ａ、レーン７，Ａ/Ｇ、レーン８，Ａ=Ｔ、レーン９，Ｇ=Ｃ、レーン１０，Ｇ/Ａ、レーン１１，Ｇ/Ｇ、レーン１２，Ｇ/Ｔ、レーン１３，Ｔ/Ｃ、レーン１４，Ｔ=Ａ、レーン１５，Ｔ/Ｇ、レーン１６，Ｔ/Ｔ。電気泳動の条件などは図１と同じ。

図6．精製した MutM のC／Cのミスマッチに対する結合

大腸菌 MutM タンパクを GST 融合タンパクとして発現させ、回収したタンパクを用いてゲルシフトアッセイを行った。それぞれのレーンのサンプルは、レーン1，大腸菌野生株の粗抽出液、レーン2，大腸菌 *mutM* 欠損株の粗抽出液、レーン、3，精製した GST-MutM タンパク、レーン4，精製した GST-MutM タンパクからトロンビンプロテアーゼ処理によって GST-tag を除去した MutM タンパク。プローブはすべてC／Cのミスマッチを含むもの。レーン4のバンドが少しネイティブの MutM タンパクよりも移動度が下がるのは、GST-tag を除去してもアミノ酸が8分子余分に結合しているためだと思われる。MutM タンパクの分子量は約 31 kDa、GST-tag は約 26 kDa。電気泳動の条件などは図1と同じ。

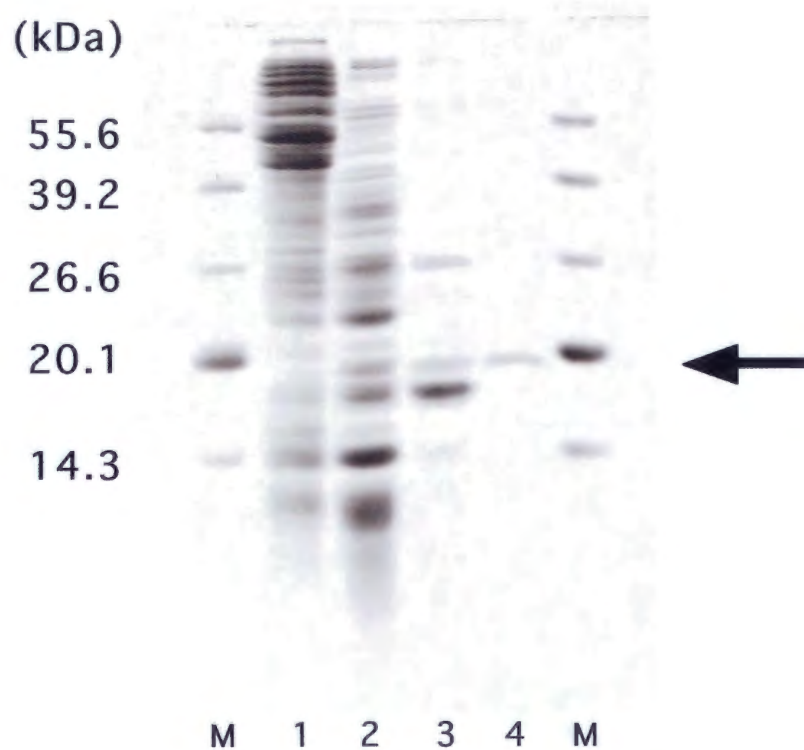



図７．二つ目のＣ/Ｃミスマッチ結合タンパクの精製。
各精製段階のサンプルを 17.5 %の SDS-PAGE で電気泳動し、クマシーブルーで染色した。各レーンのサンプルは、レーンM，分子量マーカー（上から 55.6、39.2、26.6、20.1、14.3 kDa）、レーン１，硫酸アンモニウム 40-60 %飽和分画、レーン２，HiTrap Q 陰イオン交換カラム分画、レーン３，HiTrap-Heparin カラム分画、レーン４，ゲル濾過分画。

| | |
|---|---|
| purified protein | $NH_2$-VDKRE ?YTKE- |
| | * * * * *   * * * * |
| *E. coli* FabA | $NH_2$-MVDKRESYTKEDLLASGRGELFGAKGPQLP- |

図８．精製したタンパクのＮ末端配列と大腸菌 FabA タンパクの Ｎ 末端配列の比較

精製したタンパクのアミノ酸配列を国立遺伝学研究所の大腸菌 ORF データベースで検索したところ、大腸菌 FabA タンパクのN末端の配列とほぼ一致した。図中の精製タンパクの配列の中の？はシークエンス時に解析できなかったものである。

図９．*fabA* 変異株と *mutM* 変異株のＣ／Ｃミスマッチ結合能の確認

これまでの実験の結果、Ｃ／Ｃのミスマッチを認識するタンパクが MutM と FabA とわかったので、これらの変異株から粗抽出液を取り、Ｃ／Ｃのミスマッチに結合するバンドが無くなるかどうかを確認した。各レーンの粗抽出液は、レーン１，大腸菌野生株、レーン２，*fabA* 変異株、レーン３，*mutM* 変異株。電気泳動の条件などは図１と同じ。

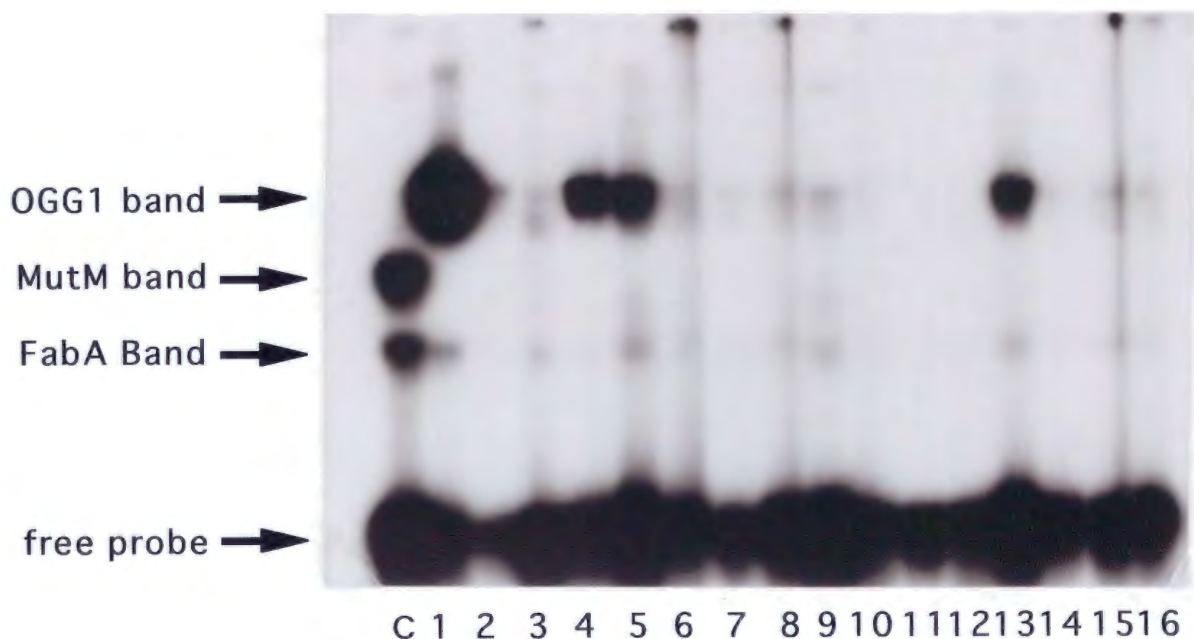



図１０．大腸菌内で発現させた出芽酵母の OGG1 タンパクのミスマッチ結合能の確認

大腸菌 MutM の機能ホモログである出芽酵母 *OGG1* 遺伝子をクローニングしたプラスミドを作成し、大腸菌 *mutM* 欠損株に導入した。この株から得られた粗抽出液を用いて、様々なミスマッチに対する結合能力の特異性を調べた。各レーンで用いられているのプローブは、レーンC，大腸菌野生株の粗抽出液にC/Cのプローブを加えたもの、レーン１，C/C、レーン２，C/A、レーン３，C=G、レーン４，C/T、レーン５，A/C、レーン６，A/A、レーン７，A/G、レーン８，A=T、レーン９，G=C、レーン１０，G/A、レーン１１，G/G、レーン１２，G/T、レーン１３，T/C、レーン１４，T=A、レーン１５，T/G、レーン１６，T/T。電気泳動の条件などは図１と同じ。

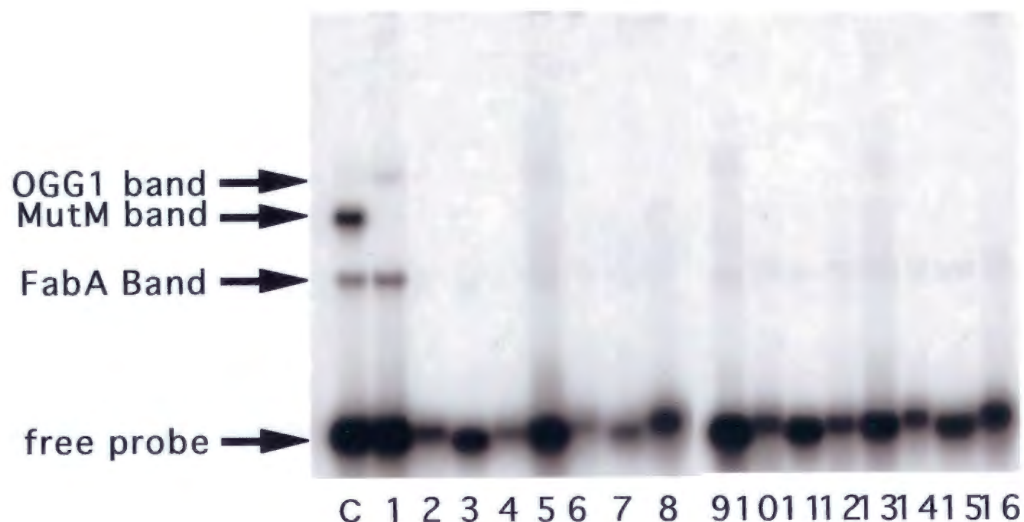



図１１．大腸菌内で発現させたヒトの OGG1 タンパクのミスマッチ結合能の確認

大腸菌 MutM の機能ホモログであるヒト *OGG1* 遺伝子をクローニングしたプラスミドを大腸菌 *mutM* 欠損株に導入した。この株から得られた粗抽出液を用いて、様々なミスマッチに対する結合能力の特異性を調べた。各レーンで用いられているプローブは、レーンＣ，大腸菌野生株の粗抽出液にＣ/Ｃのプローブを加えたもの、レーン１，Ｃ/Ｃ、レーン２，Ｃ/Ａ、レーン３，Ｃ=Ｇ、レーン４，Ｃ/Ｔ、レーン５，Ａ/Ｃ、レーン６，Ａ/Ａ、レーン７，Ａ/Ｇ、レーン８，Ａ=Ｔ、レーン９，Ｇ=Ｃ、レーン１０，Ｇ/Ａ、レーン１１，Ｇ/Ｇ、レーン１２，Ｇ/Ｔ、レーン１３，Ｔ/Ｃ、レーン１４，Ｔ=Ａ、レーン１５，Ｔ/Ｇ、レーン１６，Ｔ/Ｔ。電気泳動の条件などは図１と同じ。

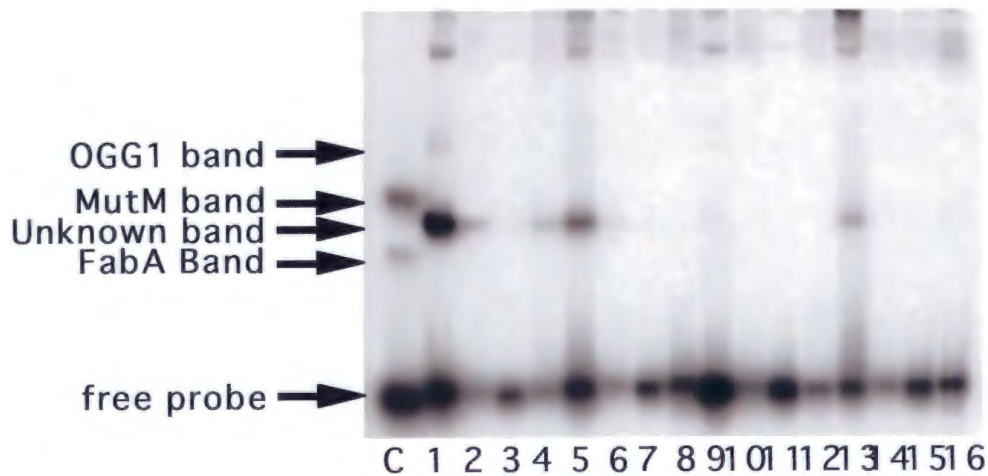



図１２．出芽酵母野生株のミスマッチ結合タンパクの確認

出芽酵母野生株から得られた粗抽出液を用いて、様々なミスマッチに対する結合能力の特異性を調べた。各レーンで用いられているプローブは、レーンC，大腸菌野生株の粗抽出液にC/Cのプローブを加えたもの、レーンY，大腸菌 *mutM* 欠損株に出芽酵母 OGG1 を発現させるプラスミドを導入した株から得た粗抽出液にC/Cのプローブを加えたもの、レーン１，C/C、レーン２，C/A、レーン３，C=G、レーン４，C/T、レーン５，A/C、レーン６，A/A、レーン７，A/G、レーン８，A=T、レーン９，G=C、レーン１０，G/A、レーン１１，G/G、レーン１２，G/T、レーン１３，T/C、レーン１４，T=A、レーン１５，T/G、レーン１６，T/T。電気泳動の条件などは図１と同じ。

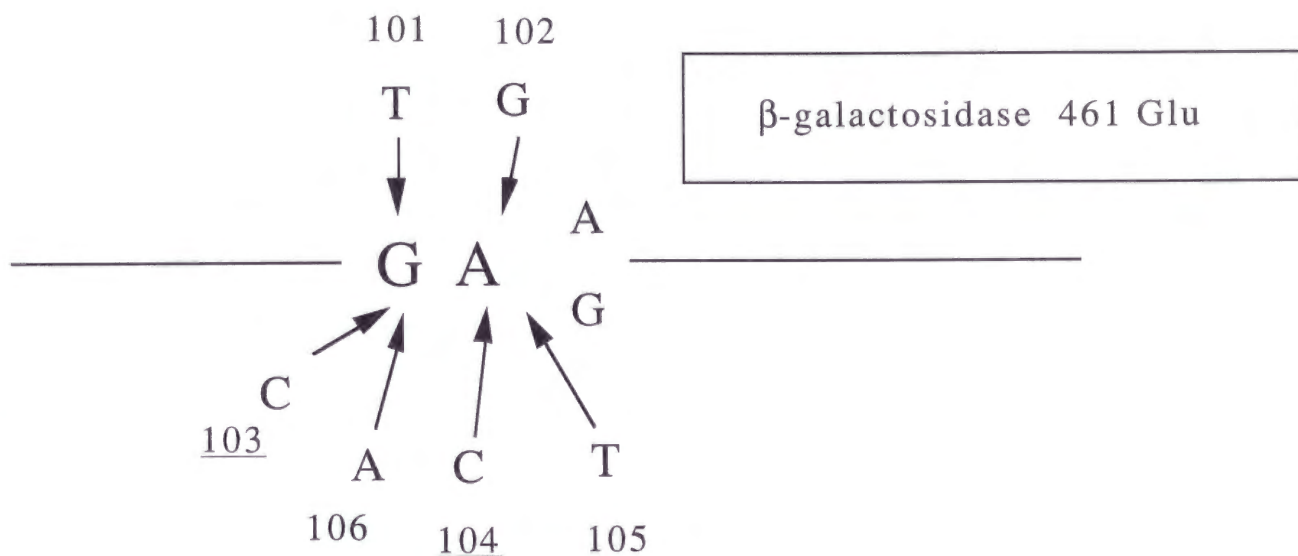



図１３．Miller の CC シリーズの概要

これらの株は、染色体上の *lacZ* 遺伝子が欠失しており、F' 因子上に *lacZ* 変異遺伝子が存在する。この遺伝子産物は、461 番目のアミノ酸が変異しており、このままでは活性がないため、ラクトース最少培地上ではコロニーを形成しないが、CC101 ～ CC106 株でそれぞれ特定の塩基置換が生じると、β-galactosidase 活性を回復するので、ラクトース最少培地上でもコロニーを形成できる。塩基置換の種類は以下の通り。CC101 : AT-CG, CC102 : GC-AT, CC103 : GC-CG, CC104 : GC-TA, CC105 : AT-TA, CC106 : AT-GC

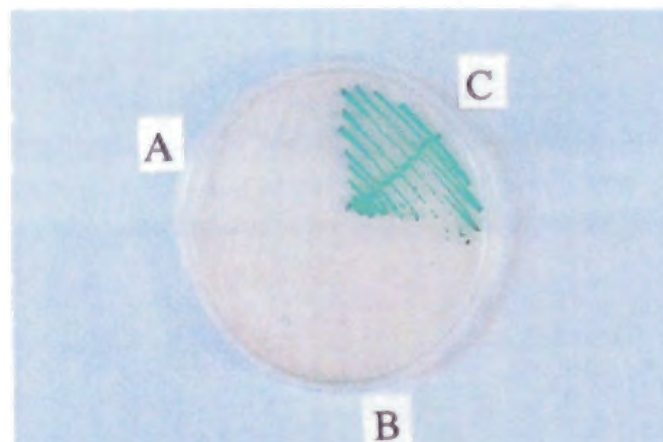



図１４．パピーレのによるミューテーター表現型の確認の例。
CC103、CC103*mutM*、CC103*mut39* の各株を X-Gal、P-Gal を含むグルコース最小培地に滅菌爪楊枝を使って拡げた。このプレートを37℃で６日培養したもの。A，CC103、B，CC103*mutM*、C，CC103*mut39*。

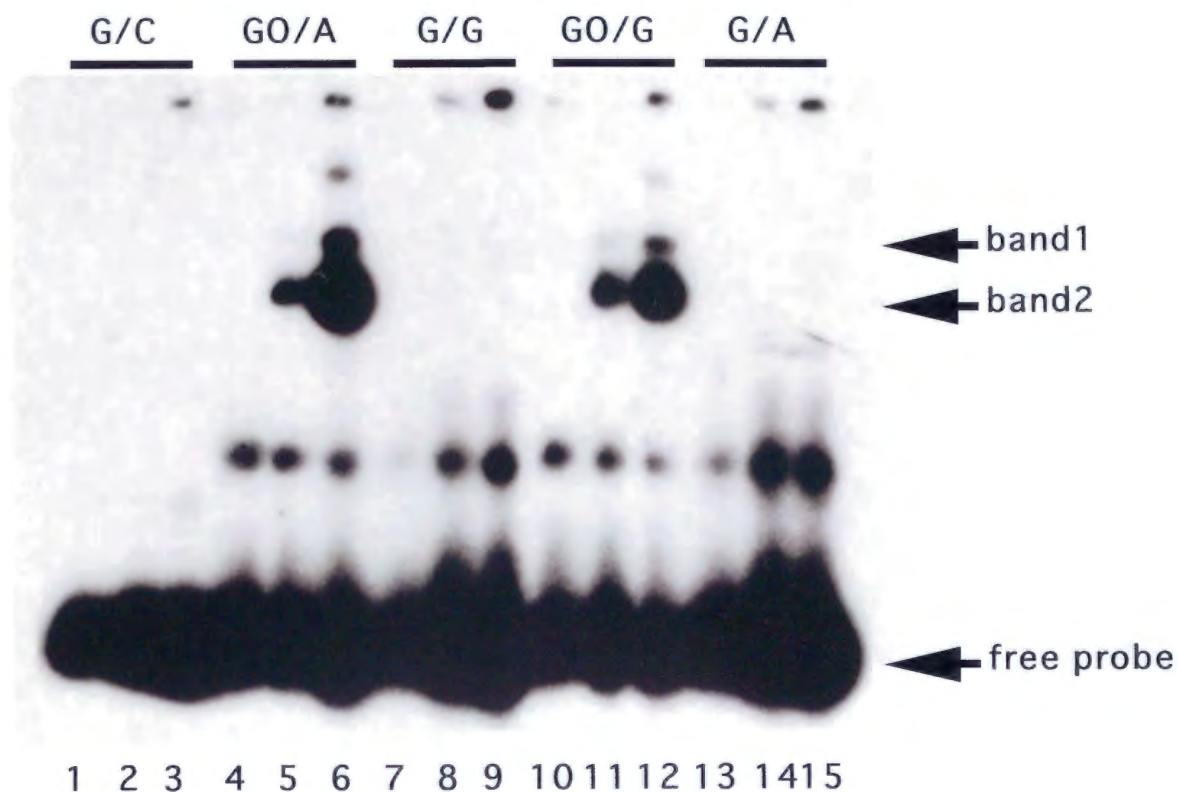



図１５．MutY タンパクの様々なミスマッチに対する結合活性の確認。

10 fmol の２本鎖ＤＮＡと MutY タンパクを反応させた。プローブは、レーン１～３，G/C。レーン４～６，GO/A。レーン７～９，G/G。レーン１０～１２，GO/G。レーン１３～１６，G/A。また、レーン１、４、７、１０、１３，MutY を加えずに反応させたもの。レーン２、５、８、１１、１４，40 ngの MutY タンパクを加えて反応させたもの。レーン３、６、９、１２，160 ng の MutY タンパクを加えて反応させたもの。

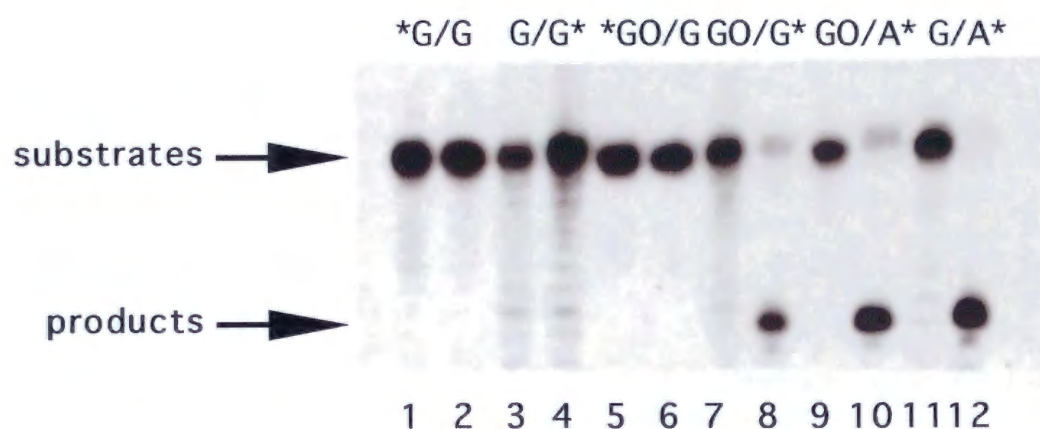



図１６．MutY タンパクによる８－オキソグアニンを含むオリゴＤＮＡの切断活性の確認。

基質ＤＮＡ 20 fmol を420 ng の MutY タンパクと３７℃で１４時間反応させた。レーン１、３、５、７、９、１１，MutY を加えず反応させたもの。レーン２、４、６、８、１０、１２，MutY を加えて反応させたもの。*のついている方のＤＮＡ鎖を標識している。

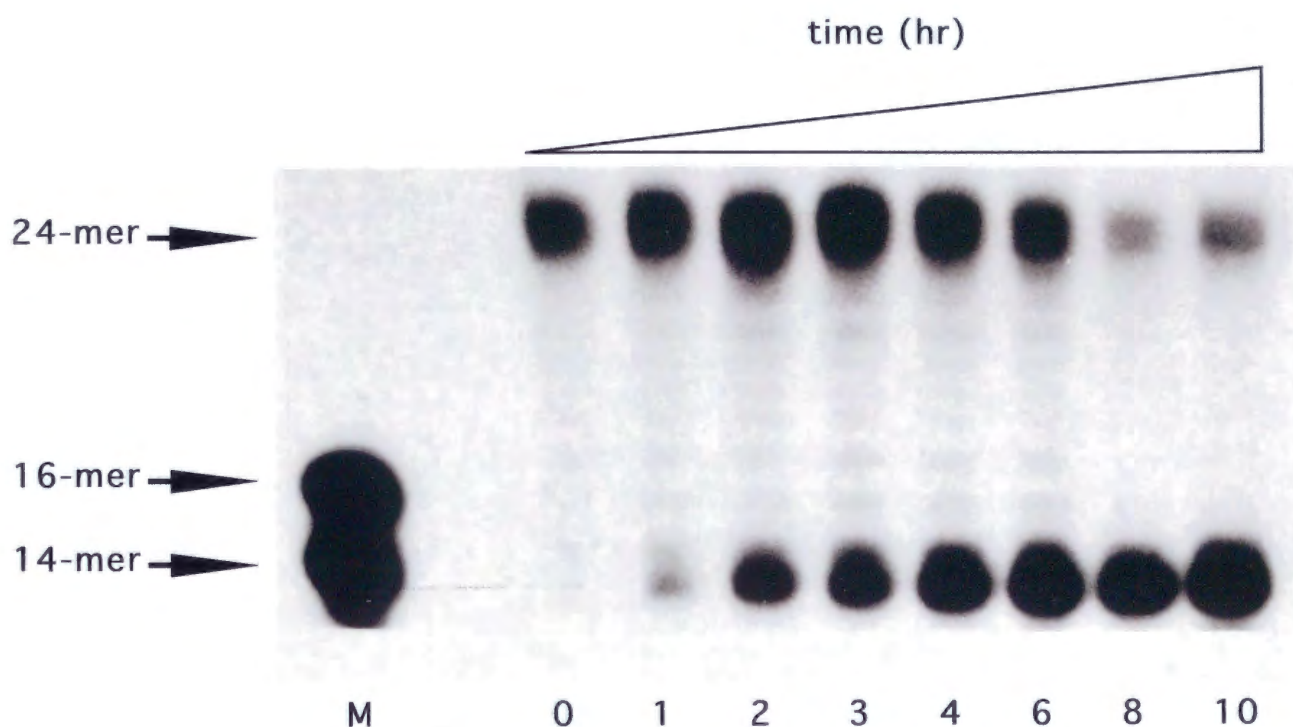



図１７．８－オキソグアニンとグアニンのミスマッチに対する MutY タンパクのグリコシラーゼ反応のタイムコース。
８－オキソグアニンとグアニンのミスマッチを含む２本鎖ＤＮＡ 20 fmol を 60 ng の MutY タンパクとともに３７℃で０～１０時間反応させた。

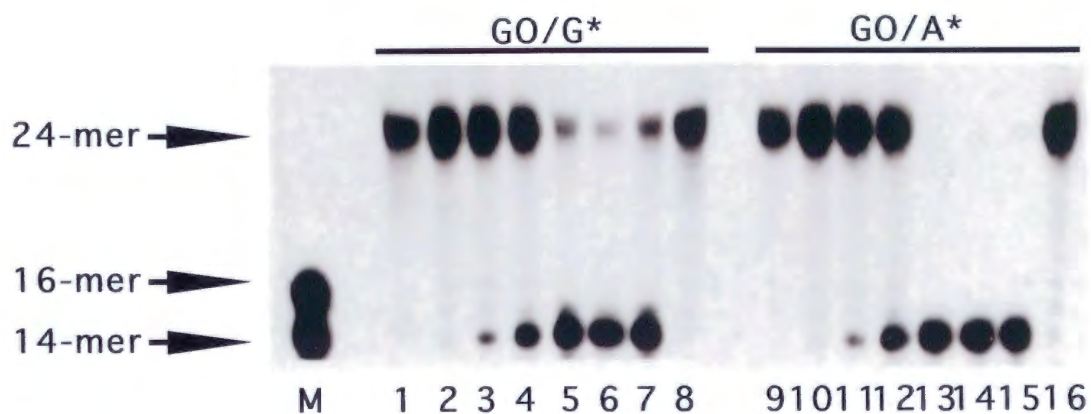



図１８． ８－オキソグアニンとグアニンやアデニンのミスマッチに対する MutY タンパクの比活性の測定。

20 fmol の８－オキソグアニンとグアニン（レーン１～８）８－オキソグアニンとアデニン（レーン９～１６）のミスマッチを含む２本鎖ＤＮＡを様々な量の MutY タンパクと反応させた。MutY タンパクの量は、レーン１、８、９、１６，0 ng、レーン２、１０，1.2 ng、レーン３、１１，2.4 ng、レーン４、１２，6 ng、レーン５、１３，30 ng、レーン６、１４，45 ng、レーン７、１５，60 ng。*のついている方のＤＮＡ鎖を標識している。

| AB1157 | wild-type |
|---|---|
| CC103 | GC-CG indicater strain |
| CC104 | GC-TA indicater strain |
| BL21 | *lon*, *ompT* |
| 21336 | *uvrD260*::Tn*5* markers of AB1157 |
| BW313 | *ung-1*, *dut-1*, *relA1*, *spoT1* |
| RPC501 | *nfo-1*::*kan*, Δ(*xth-pncA*)90 marlers of AB1157 |
| JC7623 | *recB21*, *recC22*, *sbcB15* markers of AB1157 |
| WD8014 | *mutD*, *rpsL* |
| CC104*mutM* | *mutM*::Tn*10* |
| CC104*mutMmutY* | *mutM*::Tn*10*, *mutY*::*kan* |
| SYT5 | *mutT*::*Cm* |
| CC104*mutL* | *mutL*::*kan* |
| BMH71-18 | *mutS215*::Tn*10* |
| CSH117 | *mutY*::Tn*10* |
| NJK2004 | *nth*::Cm, *nei*::kan markers of SY5 |
| GC4468*hupAB* | *hupA16*(kan), *hupB*(Cm) |
| WU3610-45 | *mfd* |
| IC3126 | Δ(*umuDC*)::Cm, *uvrA115*, *lon11* |
| MS23 | *alkA1* |
| CS101 | *recA269*::Tn*10*, *ruv-51*, *rpsL31* |
| CH1692 | *topA57*(Am), *tolC*::Tn*10*, *rpsL* |
| N3055 | *uvrA277*::Tn*10* |
| MG1655 | *dps*::kan |
| TM41 | *topB*::kan markers of AB1157 |
| RP4182 | Δ(*dcm*, *vsr*), *rpsL* |
| C218 | (*ada25 alkB*)::Cm, *ogt-1*::kan |
| GW3773 | *mutH471*::kan, *rpsL31* |

表１．本研究で使用した大腸菌株。
変異箇所については主なもののみしめした。

1: 5'- AGCTTGGCTGCAGGTCGACGGATCCCCGGGAATT- 3'

2: 5'- AGCTTGGCTGCAGGTAGACGGATCCCCGGGAATT- 3'

3: 5'- AGCTTGGCTGCAGGTGGACGGATCCCCGGGAATT- 3'

4: 5'- AGCTTGGCTGCAGGTTGACGGATCCCCGGGAATT- 3'

5: 5'- AATTCCCGGGGATCCGTCCACCTGCAGCCAAGCT- 3'

6: 5'- AATTCCCGGGGATCCGTCAACCTGCAGCCAAGCT- 3'

7: 5'- AATTCCCGGGGATCCGTCGACCTGCAGCCAAGCT- 3'

8: 5'- AATTCCCGGGGATCCGTCTACCTGCAGCCAAGCT- 3'

9: 5'- CCGGAATTCATGTCTTATAAATTCGG- 3'

10: 5'- GCCCAAGCTTCTAATCTATTTTTGCTTC- 3'

11: 5'- CCAAAATCATTAGGGGATTCATCAG- 3'

12: 5'- AACAACAGTGAATTCGGTGACCAT- 3'

13: 5'- ATATAGTCGACGTTGCAGGAAAGTA- 3'

14: 5'- GAACTAGTGCATCCCCCGGGCTGC- 3'

15: 5'- GAACTAGTGGATCCCCCGGGCTGC- 3'

16: 5'- GAACTAGTGOATCCCCCGGGCTGC- 3'

17: 5'- GCAGCCCGGGGGATGCACTAGTTC- 3'

18: 5'- GCAGCCCGGGGGATACACTAGTTC- 3'

19: 5'- GCAGCCCGGGGGATCCACTAGTTC- 3'

表２．今回の実験に用いた合成 DNA。
１～８はＣ／Ｃミスマッチなどのゲルシフトのプローブに用いた。９、１０は *yOGG1* 遺伝子のクローニングに用いた。１１は *mut39* 遺伝子のシークエンス用のプライマーとして用いた。１２、１３は *mutY* 遺伝子のクローニングに用いた。１４～１９は MutY タンパクに関するゲルシフトアッセイやニッキングアッセイのプローブとして用いた。図中の下線は相補鎖と組み合わせたときにミスマッチを形成する部分。１６の下線部「O」は８－オキソグアニンを示す。

| Group | Strain | Papillation | Mutants per $10^8$ cells | |
|---|---|---|---|---|
| | | | $Lac^+$ | $Rif^r$ |
| I | GC39 | ++ | 148 | 188 |
| II | GC9 | ++ | 125 | 1550 |
| | GC24 | ++ | 101 | 1460 |
| III | GC40 | + | 15 | 1280 |
| | GC41 | + | 28 | 2560 |
| — | CC103 | - | 3 | 11 |

表３．得られた５つの変異株のミューテーターとしての性質の解析
各変異株及び野生株のパピーレの出方、$Lac^+$への復帰変異の頻度、$Rif^r$への変異頻度を測定した。Papillationはパピーレの出方を観察したもの。
++，多くでる、+，でる、-，ほとんどでない。

| Strain | Number of $Lac^+$ revertants per $10^8$ cells |
|---|---|
| CC103 | 1 |
| CC103*mutY* | 120 |
| CC104*mutY*/pGEX-MutY | 18 |
| CC103*mutM* | 2 |
| CC103*mut39* | 146 |
| CC103*mut39*/pGEX-MutY | 13 |
| CC103*mutMmut39* | 149 |
| CC104 | 6 |
| CC104*mut39* | 232 |
| CC104*mutY* | 204 |
| CC104*mutY*/pGEX-MutY | 18 |

表４．*mutM*、*mutY*、*mut39* の変異とプラスミド pGEX-MutY の CC103 や CC104 の $Lac^+$ への復帰変異頻度に対する影響。
各株の終夜培養菌液をラクトース最小培地にまき、３７℃で CC103 株由来のものは６～１０日、CC104 株由来のものは２日培養した。また、ベクターである PGEX-4T-3 を導入しても $Lac^+$ への復帰変異頻度に影響はなかった。

添付論文

*To be revised to: Nucleic Acids Res.*

# Identification of proteins of *Escherichia coli* and *Saccharomyces cerevisiae* that specifically bind to C/C mismatches in DNA

Takehisa Nakahara, Qiu-Mei Zhang, Kazunari Hashiguchi and Shuji Yonei*

Department of Biological Sciences, Graduate School of Science, Kyoto University, Kitashirakawa-Oiwakecho, Sakyo-ku, Kyoto 606-8502, Japan

Running title: C/C mismatch-binding proteins of *E. coli* and *S. cerevisiae*

---

*Correspondence author.

Tel: +81-75-753-4097; Fax: +81-75-753-4087;

E-mail: yonei@kingyo.zool.kyoto-u.ac.jp

## ABSTRACT

The pathways leading to G:C→C:G transversions and their repair mechanisms remain uncertain. C/C and G/G mismatches arising during DNA replication are a potential source of G:C→C:G transversions. The *Escherichia coli mutHLS* mismatch repair pathway efficiently corrects G/G mismatches, whereas C/C mismatches are a poor substrate. *E. coli* must have a more specific repair pathway to correct C/C mismatches. In this study, we performed gel-shift assays to identify C/C mismatch-binding proteins in cell extracts of *E. coli.* By testing heteroduplex DNA (34-mers) containing C/C mismatches, two specific band shifts were generated in the gels. The band shifts were due to mismatch-specific binding of proteins present in the extracts. Cell extracts of a mutant strain defective in MutM protein did not produce a low-mobility complex. Purified MutM protein bound efficiently to the C/C mismatch-containing heteroduplex to produce the low-mobility complex. The second protein, which produced a high-mobility complex with the C/C mismatches, was purified to homogeneity, and the amino acid sequence revealed that this protein was the FabA protein of *E. coli*. The high-mobility complex was not formed in cell extracts of a *fabA* mutant. From these results it is possible that MutM and FabA proteins are components of repair pathways for C/C mismatches in *E. coli*. Furthermore, we found that *Saccharomyces cerevisiae* OGG1 protein, a functional homologue of *E. coli* MutM protein, could specifically bind to the C/C mismatches in DNA.

## INTRODUCTION

Base pair mismatches arise continually in DNA as a result of several mechanisms, including errors during DNA replication, spontaneous oxidation of bases and recombination between homologous but non-identical DNA sequences (1-6). The repair of such mismatched bases is essential for reducing the frequencies of spontaneous mutations and maintaining the integrity of the genome (4-10). In particular, low rates of spontaneous G:C→C:G transversions would be achieved not only by the correction of base mismatches during DNA replication but also by the prevention and removal of oxidative base damage in DNA (4,7,8,11). We previously found that MutY protein prevents the generation of G:C→C:G transversions by removing the unmodified guanine from the 8-oxoguanine:guanine mispairs in *Escherichia coli* (11).

In addition, C/C and G/G mispairs generated during DNA replication would be a source of the mutations. Bacteria and mammalian cells have evolved several repair mechanisms to deal with mismatches (5,7-10). In *E. coli*, a general repair system referred to as the *mutHLS* system corrects most replication errors with efficiencies reflecting the frequency and specificity of mispair formation (2,5,7,8). MutS and MutL are involved in the initiation step of mismatch repair. MutS recognizes and binds to single-base mismatches in DNA, and MutL forms a complex with MutS and MutH. Binding of MutL to MutH results in activation of the MutH endonuclease (2,4,5,7,8). Mutations in the *mutS*, *mutL* or *mutH* gene significantly enhance the frequencies of spontaneous mutations (4,5,7,8). However, the mutator characteristics do not result in enhancing G:C→C:G transversions (4,8,11). In *E. coli*, the *mutHLS* mismatch-repair pathway efficiently corrects G/G mismatches, but not C/C mismatches (12-14). In addition, *E. coli* MutS protein shows the lowest affinity for C/C mismatches *in vitro* (15,16). These facts suggest that *E. coli* has a more specific repair pathway to correct C/C mismatches for

preventing G:C→C:G transversions (17). Recently Fleck *et al.* (18,19) reported that there are two C/C mismatch-binding proteins in *Schizosaccharomyces pombe*. These proteins might constitute a C/C mismatch repair system. Thus, besides the MutSL-like pathway, *S. pombe* has an additional pathway of mismatch repair that corrects C/C mismatches (18,19). It is of interest to know whether or not *E. coli* has an equivalent pathway.

In this study, we performed gel-shift assays to identify and characterize C/C mismatch-binding proteins in order to look for an additional repair pathway for C/C mismatches in *E. coli*. The present experiments revealed that there are two C/C mismatch-binding proteins, MutM and FabA, in cell extracts prepared from *E. coli* cells. The binding of these proteins to C/C mismatches was independent of MutS and MutL. Furthermore, it was also demonstrated that *Saccharomyces cerevisiae* OGG1 protein, a functional homologue of *E. coli* MutM protein (20-22), could bind specifically to C/C mismatches in DNA.

## MATERIALS AND METHODS

### Bacterial strains and media

Bacterial strains used in this study were derivatives of *E. coli* K-12. The strains and their relevant genotypes were AB1157 (wild-type), 21336 (*uvrD260*::Kan), BW313 (*ung-1 dut-1*), RPC501 (*nfo-1*::Tn5), JC7623 (*recB21 recC22 sbcB15*), WD8014 (*mutD5*), CC104*mutM* (*mutM*::Tet), SYT5 (*mutT*::Cm), CC104*mutL* (*mutL*::Kan), BMH71-18 (*mutS215*::Tet), CSH117 (*mutY*::Tet), NJK2004 (*nth*::Cm *nei*::Kan), GC4468*hupAB* (*hupA16*::Kan *hupB*::Cm), WU3610-45 (*mfd*), IC3126 (Δ(*umuDC*)::Kan), MS23 (*alkA1*), CS101 (*recA269*::Tet),

CH1692 (*topA57 tolC*::Tet), N3055 (*uvrA277*::Tet), MG1655 (*dps*::Kan), TM41 (*topB*::Kan), RP4182 (Δ(*dcm vsr*)), C218 (Δ(*ada25 alkB*)::Cm *ogt-1*::Kan) and GW3773 (*mutH471*::Kan). *E. coli* CC104 was *ara* Δ(*gpt-lac*)*5 rpsL*/F' (*lacI378 lacZ461* $proA^+B^+$) (23). Bacterial cells were grown in LB medium (4) at 37°C with aeration. The final concentrations of tetracycline, kanamycin, chloramphenicol and ampicillin were 15, 50, 20 and 50 μg/ml, respectively.

**Enzymes and chemicals**

Tetracycline hydrochloride, kanamycin and chloramphenicol were purchased from Wako Pure Chemicals. Ampicillin was the product of Meiji Seika. T4 polynucleotide kinase was obtained from New England Biolabs. Glutathione-Sepharose 4B and thrombin protease were purchased from Amersham Pharmacia Biotech. Isopropyl-1-thio-β-D-galactopyranoside (IPTG), *Eco*RI, *Hind*III and KOD Plus DNA polymerase were obtained from TOYOBO. [γ-$^{32}$P] ATP (>259 TBq/mmol) was the product of ICN Biomedicals Inc.

**Construction of recombinant *S. cerevisiae OGG1* gene**

The plasmid overproducing *S. cerevisiae* OGG1 protein was constructed according to the method of Girard *et al.* (24). We synthesized two PCR primers, YOGG1-For (5'-CCGGAATTCATGTCTTATAAATTCGG-3') and YOGG1-Rev (5'-GCCCAAGCTTCTAATCTATTTTTGCTTC-3'). *S. cerevisiae* genomic library, kindly supplied by Dr. M. Ajimura (National Institute of Radiological Sciences, Japan), was used as a template for the PCR reaction to amplify the *S. cerevisiae OGG1* gene. The primer YOGG1-For was used to engineer an *Eco*RI restriction site at the beginning of *OGG1* gene. Primer

YOGG1-Rev was used to introduce a *Hind*III restriction site at the end of *OGG1* gene. The PCR was performed with a high-fidelity KOD Plus DNA polymerase, initiated by a 3 min incubation at 94°C, followed by 30 cycles of 94°C/30 sec, 50°C/30 sec and 68°C/90 sec, with a final extension of 4 min at 72°C. The 1.1 Kb-amplified fragment containing the whole coding region of the *OGG1* gene was inserted into plasmid vector pKK223-3 (Pharmacia) after digestion by *Eco*RI and *Hind*III. The resulting plasmid was named pKK223-3-YOGG1. To assure activity of protein encoded by the recombinant gene, *E. coli* CC104*mutMmutY* was transformed with the plasmid pKK223-3-YOGG1. The plasmid could suppress the frequency of spontaneous G:C→T:A transversions in *E. coli* CC104*mutMmutY*.

**Preparation of cell extracts**

Overnight cultures of *E. coli* were harvested by centrifugation, and the cell pellets (about 1 g wet weight) were resuspended in 3 ml of buffer A (50 mM Tris-HCl, pH 8.0, and 1 mM EDTA) and stored at -80°C. The samples were diluted 3-fold with buffer A and then sonicated in an ice-water bath, followed by centrifugation at 8,000 rpm for 30 min and subsequently 12,000 rpm for 30 min at 4°C.

**Gel shift assays**

The synthetic 34-mer oligonucleotides 5'-AGCTTGGCTGCAGGTXGACGGATCCCCGGGAATT-3' were $^{32}$P-labeled at the 5'-end with [γ-$^{32}$P]ATP by T4 polynucleotide kinase and then annealed with their complementary oligonucleotides 5'-AATTCCCGGGGATCCGTCYACCTGCAGCCAAGCT-3', where X and

Y represent purine (A, G) and pyrimidine (T, C) bases, respectively.

Binding of proteins to the duplex oligonucleotides was monitored by gel shift assays according to the method of Fleck *et al.* (18) with a slight modification. Protein extracts were incubated for 30 min at 4°C with 23.5 fmol of $^{32}$P-labeled oligonucleotides in 25 mM Tris-HCl (pH 8.0) containing 0.5 mM dithiothreitol, 4 mM spermidine, 0.5 mM EDTA, 10% glycerol, 25 mM NaCl, 25 mM KCl, 0.01 mM $ZnCl_2$ and 0.125 mM each dNTP. Sonicated calf thymus DNA (5 μg) was added to the reaction mixture. Electrophoresis was performed at 100 V on non-denaturing 12% polyacrylamide gels in TBE buffer (89 mM Tris-borate, pH 8.0, and 2.5 mM EDTA). The gels were dried and then autoradiographed at -80°C using Fuji RX film.

**Purification of GST-MutM proteins**

A plasmid expressing GST-MutM fusion protein was kindly supplied by Dr. J. Yokota (National Cancer Center Research Institute, Japan). The GST-*mutM* fusion gene contains the whole coding region of the *E. coli mutM* gene (amino acids 1-269) (25,26). The fusion protein was expressed in *E. coli* strain BL21 with 1 mM IPTG and purified on a glutathione-Sepharose 4B column. The GST-MutM fusion protein was further treated with thrombin protease to remove the GST-tag.

**Purification of C/C mismatch-binding protein**

The cell extracts of *E. coli mutM* mutant was divided into 4 fractions: 0-20%, 20-40%, 40-60% and 60-80% saturation of ammonium sulfate, and the binding activity to mismatch-containing oligonucleotides in each fraction was tested. The 40-60% ammonium sulfate

saturation fraction exhibited the highest binding activity. The precipitates obtained with 40-60% ammonium sulfate saturation were dissolved in buffer A and dialyzed against the same buffer (Fraction I). Fraction I was loaded on a HiTrapQ anion exchange column (Amersham Pharmacia Biotech, 5 ml x2), and the column was washed with buffer B (20 mM Tris-HCl, pH 8.0, and 150 mM NaCl). The C/C mismatch-binding activity was not retained under these conditions. The active fraction (Fraction II) was then applied to a HiTrap-Heparin column (Amersham Pharmacia Biotech, 5 ml) equilibrated with buffer C (20 mM Hepes, pH 7.6, and 150 mM NaCl) and eluted with a linear salt gradient (150-650 mM NaCl in buffer C). The active fraction eluted at 300-375 mM NaCl was concentrated and pooled (Fraction III). Fraction III was then loaded on a gel filtration column (HiPrep 16/60 Sephacryl S-200HR, Amersham Pharmacia Biotech, 5 ml) equilibrated with buffer B. The active fraction was concentrated and pooled (Fraction IV).

**SDS-polyacrylamide gel electrophoresis and amino acid sequencing**

Fraction IV was analyzed by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) on a 17.5% polyacrylamide gel, transferred to an Immobilon-PSQ membrane (Millipore) by electroblotting and then stained with Coomassie blue. The 20-KDa protein band was excised from the membrane, and the sequence of 10 amino acid from the amino terminus of the purified protein was determined by using a gas-phase amino acid sequencer.

## RESULTS

## Identification of C/C mismatch-binding proteins in *E. coli*

To detect C/C-mismatch-binding proteins in *E. coli*, we carried out gel shift assays. Cell-free extract prepared from wild-type strain AB1157 was incubated with 5'-$^{32}$P-labeled oligonucleotides (34-mers) which contained a defined single-base mismatch, and the reaction mixtures were subsequently separated on non-denaturing polyacrylamide gels. Binding of a protein to one of the 34-mers retards its migration through the gel and results in a shifted band. The results are shown in Fig. 1. We found two dense bands produced by binding to the C/C mismatch-containing oligonucleotides, and weak bands produced by binding to the C/T and C/A substrates. No shifted bands were found for non-cytosine-containing mismatches, such as G/T. The binding activity was not affected by the sequence context around the mismatches (data not shown). The C/C mismatch-binding activities disappeared upon boiling the reaction mixtures (data not shown). Thus, the two band shifts were due to mismatch-specific binding of proteins present in the extracts.

We next analyzed the competition ability of all substrates used in this study against the C/C mismatch to determine the substrate specificity of the mismatch-binding proteins. A 50-fold excess of unlabeled oligonucleotide was used as competitor. Fig. 2 shows that, under these conditions, the production of both specific shifted bands was completely abolished by specific competition with unlabeled C/C mismatch-containing oligonucleotide, but unaffected when other oligonucleotides were used as competitors. These results indicated that the C/C mismatch is the best substrate for both binding proteins.

## C/C binding activities in crude extracts of various mutants of *E. coli*

To identify genes responsible for the C/C mismatch-binding proteins, extracts of various mutants of *E. coli* were incubated with the duplex oligonucleotide containing the C/C mismatch. Production of the low-mobility complex was completely abolished in cell extracts of the *mutM* mutant (Fig. 3), indicating that the MutM protein specifically binds to the C/C mismatch-containing oligonucleotide. On the other hand, extracts from all the mutants used, including the *mutM* mutant, displayed the high-mobility band binding to the C/C mismatch-containing oligonucleotide (Fig. 3). Thus, the two complexes contained different proteins. Furthermore, the proteins in the complexes binding of two proteins was independent of the MutS and MutL proteins, because mutations of these genes did not affect the formation of the complexes.

**Binding activity of purified MutM protein to the C/C mismatches**

It was of interest to know whether or not MutM protein could directly bind to C/C mismatches. We used a GST-fusion system for MutM overexpression (25,26). Purified GST-MutM fusion protein was further treated with thrombin protease to remove the GST-tag. Purified MutM protein bound to the oligonucleotide containing C/C mismatches with the same specificity as the binding activity in wild-type cell extracts (Fig. 4). These results indicate that the MutM protein directly binds to the C/C mismatch. We therefore concluded that the low-mobility complex consists of the C/C mismatch-containing oligonucleotide and MutM protein.

**Purification and characterization of the second C/C mismatch-binding protein**

Next, we purified the second C/C mismatch-binding protein from cell extracts of *E. coli mutM*

strain by ammonium sulfate precipitation (fraction I), followed by chromatography through HiTrapQ (fraction II), HiTrap-Heparin (fraction III) and HiPrep Sephacryl-S200HR (fraction IV). Fraction IV showed homogeneity of molecular mass (about 20 KDa) (Fig. 5). The sequence of the N-terminal 10 amino acids of the purified protein was determined to be Val-Asp-Lys-Arg-Glu-?-Tyr-Thr-Lys-Glu. This matched the sequence of a predicted polypeptide starting at the second codon of *E. coli fabA* gene (27). This FabA protein should have a molecular weight of 18,969 as predicted from its amino acid sequence. This value was consistent with our findings using SDS-PAGE of fraction IV.

**C/C mismatch-binding activities in cell extracts of *fabA* mutant of *E. coli***

Extracts of an *E. coli fabA* mutant were tested to determine whether or not they contained the C/C mismatch-binding activities. The high-mobility complex was not found with the cell extracts of the *fabA* mutant, while the low-mobility complex was present (Fig. 6). These results indicated that FabA is a protein that can recognize and bind to C/C mismatches in DNA.

**C/C mismatch-binding activities in cell extracts of *E. coli* with the *S. cerevisiae OGG1* gene**

*S. cerevisiae* and mammalian cells contain a protein, OGG1, which functions as an 8-oxoguanine-DNA glycosylase/AP lyase activity (20-22,28-30) but has no obvious structural similarity to MutM (20,21). Thus, the components that repair the 8-oxoguanine are functionally, if not structurally, conserved in bacteria, yeast and mammalian cells. So, it was

of interest to examine if the *S. cerevisiae* OGG1 protein had a C/C mismatch-binding activity as well as the *E. coli* MutM protein. The plasmid pKK223-3-YOGG1 carrying the *S. cerevisiae OGG1* gene was constructed and introduced into *E. coli mutM* mutant and extracts of the mutant cells were tested to determine whether or not they contained the C/C mismatch-binding activities. The results are shown in Fig. 7. There were two dense bands produced by binding to the C/C mismatch-containing oligonucleotides, and weak bands produced by binding to the C/T and C/A substrates. The substrate specificity of yeast OGG1 protein was the same as the binding activity in *E. coli* MutM protein (Fig. 5). These results indicated that the *S. cerevisiae* OGG1 protein could also bind to the C/C mismatches in DNA.

## DISCUSSION

Errors during DNA replication are a potential source of mismatches such as G/G and C/C (4,5,7,8). Genetic recombination between homologous but not identical DNA sequences would be another cause of such mismatches (18,31,32). Mispairs in DNA are a substrate for mismatch repair pathways. Because C/C mispairs are less stable than other types of mispairs (33), they would be a good substrate for mismatch repair. However, the *mutHLS* mismatch repair pathway does not recognize C/C mismatches (4,8,11). If left unrepaired, C/C mismatches would have mutagenic consequences and should result in G:C→C:G transversions. However, spontaneous G:C→C:G transversions are rare events (4,8,11). Therefore, another specific repair pathway must operate on C/C mismatches in *E. coli*. The present experiments revealed the formation of two mismatch-dependent DNA-protein complexes upon incubation of oligonucleotide heteroduplexes with crude extracts of *E. coli*

wild-type cells (Fig. 1). Both complexes were formed preferentially with C/C mismatch-containing oligonucleotides (Fig. 1). In addition, unlabeled duplex oligonucleotides containing C/C mismatches efficiently prevented the formation of C/T and C/A mismatch-protein complexes (Fig. 2). These results indicated that *E. coli* has at least two distinct proteins that specifically bind to C/C mismatches.

The present experiments revealed that *E. coli* MutM protein has the ability to specifically bind to the C/C mismatches. This conclusion was reached based on the following facts: gel shift assays showed that the low-mobility complex was completely absent in the cell extracts prepared from *E. coli mutM* mutant (Fig. 3). Furthermore, purified MutM protein could bind directly to the C/C mismatches (Fig. 4). MutM protein shows broad substrate specificity: formamidopyrimidine (Fapy), 8-oxoguanine, 5-hydroxycytosine and thymine glycols (34-37). The MutM protein removes these products of oxidative base damage from DNA by its DNA glycosylase activity and incises the DNA by its associated AP lyase activity (6,9,10,34-37). In this study, purified MutM protein did not incise the C/C mismatch-containing oligonucleotide (data not shown). Therefore, it is suggested that the MutM protein can recognize and bind to C/C mismatches in DNA, and requires another protein(s) to incise DNA for the repair of the mismatches. The MutM protein might be a component of a multienzymic pathway involved in C/C mismatch repair.

The present experiments, moreover, showed that the second protein that specifically binds to the C/C mismatch is the FabA protein of *E. coli*. The purified FabA protein could bind the mismatch-containing oligonucleotide to form the low-mobility complex (Figs. 5 and 6). In addition, cell extracts of the *fabA* mutant did not display the binding activity (Fig. 6). The FabA protein is a component of the fatty acid biosynthesis system in *E. coli*, and defined as β-hydroxydecanoyl-acyl carrier protein dehydrase, which introduces a double bond into a

growing fatty acid chain (27,38-40). However, no DNA-binding domain is found in the amino acid sequence of the FabA protein (27). Interestingly, *S. pombe* fatty acid synthase, p190/210 complex, binds to single-stranded and double-stranded DNA leading to condensation of large DNA aggregates (41). Furthermore, the protein is capable of renaturing complementary single-stranded DNA and exhibits strand exchange activity (41). Thus, some enzymes, which catalyze fatty acid synthesis also, have DNA-processing activities. The mechanisms of the involvement of the MutM and FabA proteins in C/C mismatch repair in *E. coli* cells are under investigation in our laboratory.

In *S. pombe*, there are at least two different mechanisms for mismatch repair pathways. The major system recognizes all base mismatches except C/C, probably like the *mutHLS* system in *E. coli*. The minor system can recognize C/C mismatches and leads to short excision tracts (42,43). Furthermore, recent reports have shown that *S. pombe* has two distinct C/C mismatch-binding proteins, termed Cmb1 and Cmb2 (19). Purified Cmb1 protein has a molecular mass of 22 KDa. The deduced amino acid sequence contains a high mobility group (HMG) domain, a motif common to a heterogeneous family of DNA-binding proteins. Unlike the Cmb1 protein, MutM and FabA do not have the HMG domain (39,44). The Cmb1 protein recognizes C/T and C/A mismatches, and, in addition, T/T mismatch, C/Δ loop, $O^6$-methylguanine:C mispair and 1,2 GpG intrastrand crosslinks produced by cisplatin (19). On the other hand, Cmb2 protein recognizes only the C/C mismatch (19). It is also suggested that components of the nucleotide excision repair system are involved in C/C mismatch repair systems (45). Thus, it is expected that DNA repair systems other than the *mutHLS* system must play a role in C/C mismatch repair in *E. coli* cells. It is possible that MutM and FabA are components of such repair pathways as mismatch recognition proteins like MutS.

In *S. cerevisiae*, rat and human cells, OGG1 proteins are identified as an 8-oxoguanine-

DNA glycosylase/AP lyase (20-22). They remove 8-oxoG paired with cytosine and guanine, like *E. coli* MutM protein (6,9,10,20-22). The present experiments demonstrated that *S. cerevisiae* OGG1 protein could specifically bind to C/C mismatches like *E. coli* MutM protein (Fig. 7). OGG1 has no significant structural homology to the *E. coli* MutM protein in the amino acid sequences (20,21). It is sure that these proteins recognize and bind to the mismatches as an essential function.

## ACKNOWLEDGEMENTS

We wish to express our thanks to Drs. J. Yokota (National Cancer Center Research Institute, Japan), M. Ajimura (National Institute of Radiological Sciences, Japan), K. Yamamoto (Tohoku University, Japan), B. Weiss (Michigan University, USA), A. Nishimura (National Institute of Genetics, Japan), M. M. Wu (Harvard University, USA), A. Holmgren (Karolinska Institute, Sweden) and B. Bachmann (Yale University, USA) for kindly supplying *E. coli* strains and plasmids. This study was supported in part by grants from the Ministry of Education, Science, Sports and Culture of Japan and the Nissan Science Foundation (to Q.-M.Z.).

## REFERENCES

1. Drake, J. W. (1991) *Annu. Rev. Genet.*, **25**, 125-146.
2. Modrich, P. (1991) *Annu. Rev. Genet.*, **25**, 229-253.
3. Smith, K. C. (1992) *Mutat. Res.*, **277**, 139-162.
4. Miller, J. H. (1992) *A Short Course in Bacterial Genetics,* Cold Spring Harbor Laboratory Press, New York.
5. Modrich, P. and Lahue, R. (1996) *Annu. Rev. Biochem.*, **62**, 101-133.
6. Wallace, S. S. (1997) *Oxidative Stress and the Molecular Biology of Antioxidant Defenses* (ed. J. G. Scandalios), Cold Spring Harbor Laboratory Press, New York, pp. 49-90.
7. Miller, J. H. (1996) *Annu. Rev. Microbiol.*, **50**, 625-643.
8. Miller, J. H. (1998) *Mutat. Res.*, **409**, 99-106.
9. Mol, C. D., Parikh, S. S., Putnam, C. D., Lo, T. P. and Tainer, J. A. (1999) *Annu Rev. Biophys. Biomol. Struct.*, **28**, 101-128.
10. Eisen, J. A. and Hanawalt, P. C. (1999) *Mutat. Res.*, **435**, 171-213.
11. Zhang, Q.-M., Ishikawa, N., Nakahara, T. and Yonei, S. (1998) *Nucleic Acids Res.*, **26**, 4669-4675.
12. Kramer, B., Kramer, W. and Fritz, H.-J. (1984) *Cell*, **38**, 879-887.
13. Su, S.-S., Lahue, R. S., Au, K. G. and Modrich, P. (1988) *J. Biol. Chem.*, **263**, 6829-6835.
14. Schaaper, R. M. (1993) *J. Biol. Chem.*, **268**, 23762-23765.
15. Wagner, R., Debbie, P. and Radman, M. (1995) *Nucleic Acids Res.*, **23**, 3944-3948.
16. Babic, I., Andrew, S. E. and Jirik, F. R. (1996) *Mutat. Res.*, **372**, 87-96.
17. Radicella, J. P., Clark, E. A. and Fox, M. S. (1988) *Proc. Natl. Acad. Sci. USA.*, **85**, 9674-

9678.

18. Fleck, O., Schar, P. and Kohli, J. (1994) *Nucleic Acids Res.*, **22**, 5289-5295.

19. Fleck, O., Kunz, C., Rudolph, C. and Kohli, J. (1998) *J. Biol. Chem.*, **273**, 30398-30405.

20. van der Kemp, P. A., Thomas, D., Barbey, R., de Oliveira, R. and Boiteux, S. (1996) *Proc. Natl. Acad. Sci. USA*, **93**, 5197-5202.

21. Nash, H. M., Bruner, S. D., Scharer, O. D., Kawate, T., Addona, T. A., Spooner, E., Lane, W. S. and Verdine, G. L. (1996) *Curr. Biol.*, **6**, 968-980.

22. Sandigursky, M., Yacoub, A., Kelley, M. R., Xu, Y., Franklin, W. A. and Deutsch, W. A. (1997) *Nucleic Acids Res.*, **25**, 4557-4561.

23. Cupples, C. G. and Miller, J. H. (1989) *Proc. Natl. Acad. Sci. USA*, **86**, 5345-5349.

24. Girard, P. M., Guibourt, N. and Boiteux, S. (1997) *Nucleic Acids Res.*, **25**, 3204-3211.

25. Nagashima, M., Sasaki, A., Morishita, K., Takenoshita, S., Nagamachi, Y., Kasai, H. and Yokota, J. (1997) *Mutat. Res.*, **383**, 49-59.

26. Shinmura, K., Kasai, H., Sasaki, A., Sugimura, H. and Yokota, J. (1997) *Mutat. Res.*, **385**, 75-82.

27. Cronan, J. E., Jr, Li, W.-B., Coleman, R., Narasimhan, M., de Mendoza, D. and Schwab, J. M. (1988) *J. Biol. Chem.*, **263**, 4641-4646.

28. Bjørås, M., Luna, L., Johnson, B., Hoff, E., Haug, T., Rognes, T. and Seeberg, E. (1997) *EMBO J.*, **16**, 6314-6322,

29. Dherin, C., Radicella, J. P., Dizdaroglu, M. and Boiteux, S. (1999) *Nucleic Acids Res.*, **27**, 4001-4007.

30. Ishida, T., Hippo, Y., Nakahori, Y., Matsushita, I., Kodama, T., Nishimura, S. and Aburatani, H. (1999) *Cytogenet. Cell Genet.*, **85**, 232-236.

31. Fox, M. S., Radicella, J. P. and Yamamoto, K. (1994) *Experientia*, **50**, 253-260.

32. Worth, L., Jr., Clark, S., Radman, M. and Modrich, P. (1994) *Proc. Natl. Acad. Sci. USA.*, **91**, 3238-3241.

33. Aboul-ela, E., Koh, F. D. and Tinoco, I., Jr. (1985) *Nucleic Acids Res.*, **13**, 4811-4824.

34. Hatahet, Z., Kow, Y. W., Purmal, A. A., Cunningham, R. P. and Wallace, S. S. (1994) *J. Biol. Chem.*, **269**, 18814-18820.

35. Tchou, J., Kasai, H., Shibutani, S., Chung, M.-H., Laval, J., Grollman, A. P. and Nishimura, S. (1991) *Proc. Natl. Acad. Sci. USA.*, **88**, 4690-4694.

36. Tchou, J., Bodepudi, V., Shibutani, S., Antoshechkin, I., Miller, J., Grollman, A. P. and Johnson, F. (1994) *J. Biol. Chem.*, **269**, 15318-15324.

37. David, S. S. and Williams, S. D. (1998) *Chem. Rev.*, **98**, 1221-1261.

38. Silbert, D. F. and Vagelos, P. R. (1967) *Proc. Natl. Acad. Sci. USA*, **58**, 1579-1586.

39. Magnuson, K., Jackowski, S., Rock, C. O. and Cronan, J. E., Jr. (1993) *Microbiol. Rev.*, **57**, 522-542.

40. Saito, K., Hamajima, A., Ohkuma, M., Murakoshi, I., Ohmori, S., Kawaguchi, A., Teeri, T. H. and Cronan, J. E., Jr. (1995) *Transgenic Res.*, **4**, 60-69.

41. Käslin, E. and Heyer, W.-D. (1994) *J. Biol. Chem.*, **269**, 14103-14110.

42. Schär, P. and Kohli, J. (1993) *Genetics*, **133**, 825-835.

43. Schär, P., Munz, P. and Kohli, J. (1993) *Genetics*, **133**, 815-824.

44. Boiteux, S., O'Connor, T. R. and Laval, J. (1987) *EMBO J.*, **6**, 3177-3183.

45. Fleck, O., Lemann, E., Schär, P. and Kohli, J. (1999) *Nature Genet.*, **21**, 314-317.

## FIGURE LEGENDS

### Fig. 1. Gel shift assays for mismatch-binding proteins

$^{32}$P-labeled double-stranded oligonucleotides containing all possible single-base mismatches (23.5 fmol) were incubated with cell extracts prepared from wild-type *E. coli*. After incubation for 30 min at 4°C, the reaction mixtures were separated by an electrophoresis on 12% non-denaturing polyacrylamide gels in TBE buffer at 100V. The gels were dried and then autoradiographed using Fuji RX films at -80°C. Lane 1, C/C; lane 2, C/A; lane 3, C=G; lane 4, C/T; lane 5, A/C; lane 6, A/A; lane 7, A/G; lane 8, A=T; lane 9, G=C; lane 10, G/A; lane 11, G/G; lane 12, G/T; lane 13, T/C; lane 14, T=A; lane 15, T/G; lane 16, T/T.

### Fig. 2. Mismatch specificity of the C/C binding activities determined by competition with different heteroduplexes

Reactions were performed with radiolabeled oligonucleotides containing the C/C mismatch and a 50-fold excess of unlabeled oligonucleotide as competitor. Lane C; without competitor; lane 1, C/C; lane 2, T/C; lane 3, A/C; lane 4, C/T; lane 5, C/A; lane 6, C=G; lane 7, A/A; lane 8, A/G; lane 9, A=T; lane 10, G=C; lane 11, G/A; lane 12, G/G; lane 13, G/T; lane 14, T=A; lane 15, T/G; lane 16, T/T.

### Fig. 3. Cytosine-containing mismatch-binding activities in cell extracts of various mutants of *E. coli*

Gel shift assays were performed as described in Fig. 1. Lane 1, *mutD*; lane 2, *mutL*; lane 3, *mutM*; lane 4, *mutT*; lane 5, *mutY*; lane 6, *ung*; lane 7, *vsr*; lane 8, *uvrD*; lane 9, *ada* and *ogt*;

lane 10, *xth* and *nfo*; lane 11, *alkA*; lane 12, *hupA* and *hupB*; lane 13, *dps*; lane 14, *mfd*; lane 15, *nth* and *nei*; lane 16, *recA*; lane 17, *mutH*; lane 18, *mutS*; lane 19, *recB*, *recC* and *sbcB*; lane 20, *topA*; lane 21, *umuC* and *umuD*; lane 22, *uvrA*; lane 23, *topB*.

**Fig. 4. Binding of purified MutM protein to various mismatches**

The GST-MutM fusion protein was expressed in *E. coli* strain BL21 with 1 mM IPTG and purified on glutathione-Sepharose 4B column. The GST-MutM fusion protein was further treated with thrombin protease to remove the GST-tag. Gel shift assays were performed as described in Fig. 1. Lane 1, C/C; lane 2, C/A; lane 3, C=G; lane 4, C/T; lane 5, A/C; lane 6, A/A; lane 7, A/G; lane 8, A=T; lane 9, G=C; lane 10, G/A; lane 11, G/G; lane 12, G/T; lane 13, T/C; lane 14, T=A; lane 15, T/G, lane 16, T/T. Lanes W and M, cell extracts prepared from wild-type and *mutM* mutant, respectively.

**Fig. 5. Purification of the second C/C binding protein**

Proteins were separated by 17.5% SDS-PAGE and stained with Coomassie Blue. Lane M; molecular weight markers (55.6, 39.2, 26.6, 20.1 and 14.3 KDa), lane 1; 40-60% ammonium sulfate saturation fraction; lane 2, active fraction after HiTrapQ anion exchange column chromatography; lane 3, active fraction after HiTrap-Heparin column chromatography; lane 4, active fraction after gel filtration column chromatography.

**Fig. 6. The C/C binding activity in cell extracts of *fabA* and *mutM* mutants**

Cell extracts of *E. coli* wild-type (lane 1) , *fabA* (lane 2) and *mutM* (lane 3) were incubated with the C/C mismatch-containing heteroduplex. Gel shift assays were performed as described in

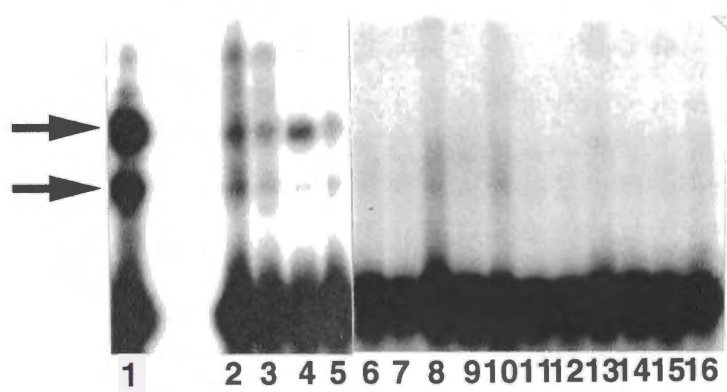
Fig. 4.

**Fig. 7. The C/C binding activity in cell extracts of *E. coli mutM* with pKK223-3-YOGG1 carrying *S. cerevisiae OGG1* gene**

$^{32}$P-labeled double-stranded oligonucleotides containing all possible single-base mismatches (23.5 fmol) were incubated with cell extract prepared from *E. coli mutM* mutant with pKK223-3-YOGG1 carrying *S. cerevisiae OGG1* gene. Gel shift assays were performed as described in Fig. 1. Lane 1, extracts from *E. coli* wild-type; lanes 2-17, extracts from *E. coli mutM* mutant with pKK223-3-YOGG1. Lanes 1 and 2, C/C; lane 3, C/A; lane 4, C=G; lane 5, C/T; lane 6, A/C; lane 7, A/A; lane 8, A/G; lane 9, A=T; lane 10, G=C; lane 11, G/A; lane 12, G/G; lane 13, G/T; lane 14, T/C; lane 15, T=A; lane 16, T/G; lane 17, T/T.

Fig.1

Nakahara *et al.*

Fig.2

Nakahara *et al.*

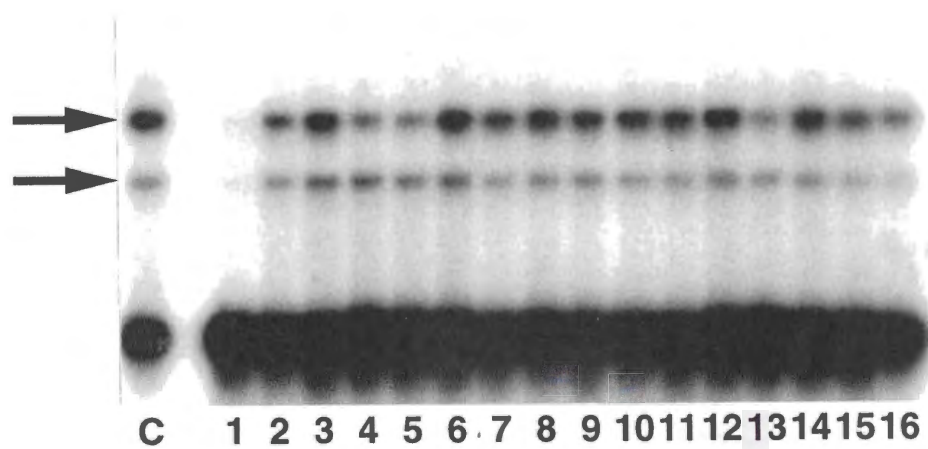



Fig.3

Nakahara *et al*

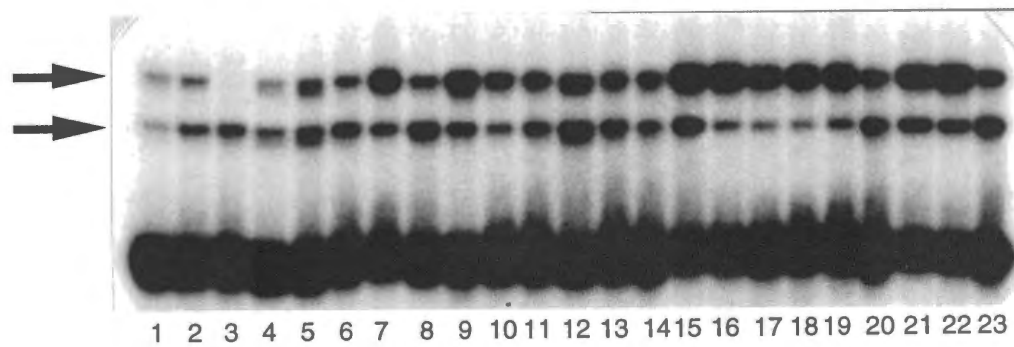
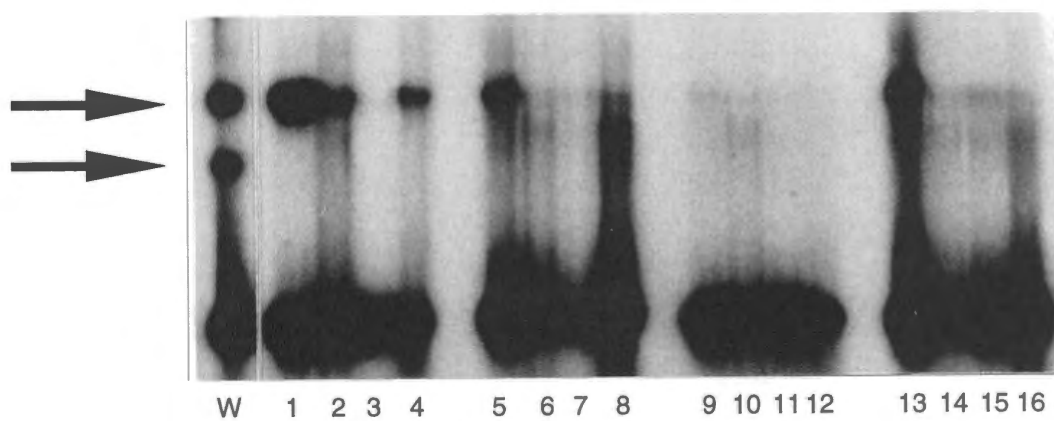



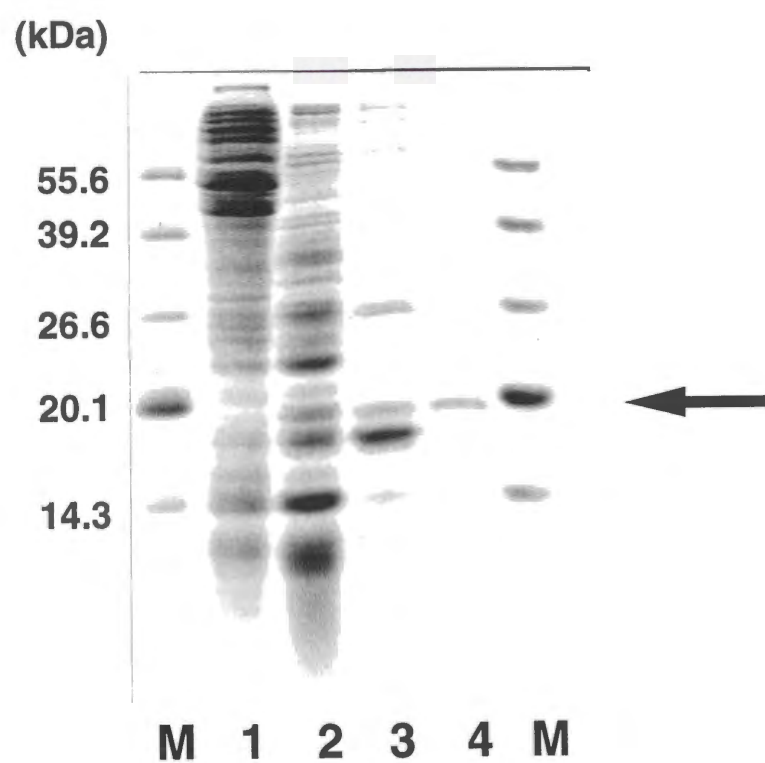
Fig.4

Nakahara *et al.*

Fig.5

Nakahara *et al.*

Fig.6

Nakahara *et al.*

Fig.7

Nakahara *et al.*